

THIS PAGE BLANK (USPTO)

A:Accession: I37005
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:L37050; NID:9606991; PIDN:AAA65624.1; PID:9606992

alignment_scores:
 Quality: 1975.00 Length: 417
 Ratio: 4.889 Gaps: 0
 Percent Similarity: 96.882 Percent Identity: 93.046

alignment_block:
 US-09-600-714-41 x I37005

Align seg 1/1 to: I37005 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCGCGGCTGTCTCCGCGCTGCGCCCTCTGGGC 50
|||||
1 MetSerSerIysTyrProArgSerValArgArgCysLeuProLeuCysAl 17
51 CCTAACCTGGAAGCAGCTCTCATCTCTCTTATTTTTCACCACT 100
|||||
17 AleuThrLeuGluAlaAlaLeuIleLeuPhePhePheThrGlnT 34
101 ATGACGCTTCCTTAGAGAGATCAAAAGGGGCTCGGCACTCATCAAGT 150
|||||
34 TyrAlaSerLeuGluAspGlnGlyLeuValAlaSerTyrGluVal 50
151 GGCCAAGATCTGACCGTGATGGCGCCATTTGGCTGGCTTCCACCTC 200
|||||
51 GlyIleAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSe 67
201 GAGTTCGCGAGACACAGCTGAGACAGTGTGGCTTCACCTCTTCATGC 250
|||||
67 rSerPheArgArgHisSerTrpSerSerValAlaPheSerLeuPheMetL 84
251 TGGCGCTTGTGTGCACTGGGCAATCCTGTGAGCGGCTTCCTGAGCCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTCTGGGAAGGTGTCATCACACACTGTTCAGTATCGGCTGGCCAC 350
|||||
101 PheProGlyGlyValValIleThrLeuPheSerIleAlaGlyLeuAlaThr 117
351 CATGAGTGTCTTGTGCTGATCTCATGATGATGATGATGATGATGATG 400
|||||
117 rThrSerAlaLeuSerValIleLeuIleSerValAlaSerValLeuGlyLysV 134
401 TCAACTGGCGCAGCTGTGTGTGATGTGCTGTGAGAGTGACAGCTTTA 450
|||||
134 AlaAsnLeuValGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCAACTGAGGATGTCATCATATATCTTTCACACACACTCCACAT 500
|||||
151 GlyThrValArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATATGACATCTACGTCGTCGACGCTATTTGGGCTGCTGTGG 550
|||||
167 ThrLeuMetHisIleTyrValIleAlaAlaIleTyrPheGlyLeuSerValA 184
551 CCTGTGCTGCCCAAGCCTCTTACCGAGGAGGAGGAGGAGATTAAGATCAG 600
|||||
184 IATrPcysLeuProLysProLeuProLysGlyThrGluAspGlyAspGln 200
601 ACAGACAGATACCCAGTTGTCTGCAATGCTGGGCGCCCTCTTCTGTG 650
|||||
201 IleAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTCTGGCCAGTTTCACTGCTGCTGCTGAGAAATGCAATCGAAA 700
|||||
217 pMetPheTrpProSerPheAsnSerAlaIleLeuArgSerProIleGluA 234
```

```

701 GGAAGAAATGCCGTTCACACACTACTATGCTGTAGCAGTACAGCTGTG 750
|||||
234 rGlyAsnAlaValAlaPheAsnThrTyrThrAlaValAlaValSerValAl 250
751 ACAGCACTTCAGGATACCTTGGCTGACCCCAAGGAGGAGATTCAGCAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerMe 267
801 GACTATATGTCACAGTCGCGGTGTGGCAGGAGGCGTGGCTGTGGTACCT 850
|||||
267 tSerTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValGlyThrS 284
851 CGTGTACACTGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||||
284 eCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATTCGCGGGGAGGCAAGTACTGCGCGGGGTGTGTAA 950
|||||
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysAs 317
951 CCGAGTCTGGGATTCGCCACAGCTCCATCATGAGGCTACACTTACAGT 1000
|||||
317 nArgValLeuGlyIleProHisSerSerValMetGlySerAsnPheSer 334
1001 TCGTGGCTGCTTGGAGAGATCATCTCATTTGCTGCTGCTGCTGAT 1050
|||||
334 rPheGlyLeuLeuGlyGluIleIleTyrIleValLeuValAlaArgHis 350
1051 ACCGTGGAGCCGCGCAATGGCATGATGGCTTCAGGCTCCCTCAGCAT 1100
|||||
351 ThrIleTrpAsnGlyAsnGlyMetIleGlyPheGlnAlaLeuLeuArgI 367
1101 TGGGAACTACAGCTTGGCCATGCTGATAGCTCTACAGTGTGCTCTGTA 1150
|||||
367 eGlyGlnPheSerLeuAlaThrThrIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTGTCTCTTAATCTTAATATGAAAGACACTCATGAGGCTAAA 1200
|||||
384 hrGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGluAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGCTGTGGAT 1250
|||||
401 TyrPheAspAspGlnValAlaPheTrpLysPheProHisLeuAlaValGly 417
1251 T 1251
417 e 417
```

seq_name: pI37076

seq_documentation_block:

Rhesus-like protein - gorilla
 C:Species: gorilla gorilla (gorilla)
 C:Date: 04-Oct-1996 #sequence-revision 04-Oct-1996 #text-change 21-Jul-2000
 C:Accession: I37076
 R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
 Biochem. Genet. 32, 201-221, 1994
 A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A:Reference number: I37003; MUID:95085595
 A:Accession: I37076
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:L37053; NID:9607011; PIDN:AAA65627.1; PID:9607012

alignment_scores:
 Quality: 1953.00 Length: 417
 Ratio: 4.870 Gaps: 0
 Percent Similarity: 96.163 Percent Identity: 92.806

alignment_block:
 US-09-600-714-41 x I37076

THIS PAGE BLANK (USPTO)

Align seg 1/1 to: 137076 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCCGGGTCTGTCCGGCGCTGCTCCCTTGAGC 50
1 MetSerSerLysTyrProArgSerValAlaGlyCysLeuProLeuGly 17
51 CCTAACATGGAAGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCT 100
17 AleuThrLeuGlnAlaAlaLeuThrLeuPheThrPhePheThr 34
101 ATGAGCTCTCTTAGAGATCAAAAGGGCTGTCGATCATCTATCAAGT 150
34 TyrAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAGATCTGACGCTGATGCGGGCATTTGGCTGGCTCTCCACCTC 200
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThr 67
201 GAGTTTCCGAGACACAGCTGGAGAGTGTGGCTTCAACCTTCATGCG 250
67 rSerPheArgGlyHisSerTyrSerSerValAlaPheAsnLeuPheMet 84
251 TGGCGCTTGCTGTCAGTGGGCAATCTGTCGAGCGCTTCTCCAGCCAG 300
84 euaLeuGlnValGlnTrrPalaIleLeuLeuAspLysPheLeuSerGln 100
301 TTCCCTCTGGGAAGGTGTCATCACACTGTTTCAGTATTCGGCTGGCC 350
101 PheProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTCTTTGCTGCTGCTGATCTCAGTGATGCTGCTTGGGAGAG 400
117 rMetSerAlaLeuSerValLeuIleSerAlaGlyAlaValLeuGlyTyr 134
401 TCAACTGGGGCAGTTGGTGTGATGGTGTGCTGGAGTGAACAGCTTGA 450
134 alAsnLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCTGAGAGTGTCTCATCAATATCTTCAACACAGACTACACAT 500
151 GlyThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACTGCTGTCGAGCCTATTTGGCTGCTGCTGG 550
167 tAsnMetMetHisIleTyrValPheAlaAlaCysPheGlyLeuSerVal 184
551 CCGTGTGCTGGCCAAAGCTCTTACCGAGGAGGAGGAGATTAAGATCAG 600
184 latrPcyLeuProLysProLeuAlaLysGlyThrGlnAspLysAspGln 200
601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGCGCCCTCTTCTGTG 650
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlnAlaLeuPheLeuTr 217
651 GATGTTCTGCCAAGTTTCAACTGCTGCTGCTGAGAAGTCCATTCGAA 700
217 pMetPheTrpProSerPheAsnSerAlaLeuLeuArgSerProIleGln 234
701 GGAGATGCGGCTGTCACACCTACTATGCTGTAGCAGTACAGCGTGTG 750
234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerVal 250
751 ACAGCATCTCAGGCTCATCTTGGCTCACCCCAAGGAAGATCAGCAA 800
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsnMe 267
801 GACTTATGTGACAGTGGCGTGTGTGGCAGAGAGCGCTGGCTGGGTACT 850
267 tThrTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValAlaThr 284
851 CGTGCACCTGATCCCTTCTCCGCTTGGCATGCTGCTGGCTGTGTG 900

```

seq_name: p1r2:137004

seq_documentation_block:

Rhesus-like protein - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
 C:Accession: 137004
 R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, D.; Socha, W.W.; Collin, Y.; Ruffi
 Blochem. Genet. 32, 201-221, 1994
 A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A:Reference number: 137003; MUID: 95085595
 A:Accession: 137004
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-417 <RES>
 A:Cross-references: GB:137049; NID:9606989; PIDN:AA65623.1; PID:9606990

alignment_scores:

Quality	Length
1952.00	417
Ratio: 4.868	Gaps: 0
Percent Similarity: 96.163	Percent Identity: 92.566

alignment_block:

US-09-600-714-41 x 137004

Align seg 1/1 to: 137004 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCCGGGTCTGTCCGGCGCTGCTCCCTTGAGC 50
1 MetSerSerLysTyrProArgSerValAlaGlyCysLeuProLeuGly 17
51 CCTAACATGGAAGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCT 100
17 AleuThrLeuGlnAlaAlaLeuThrLeuPheThrPhePheThr 34
101 ATGAGCTCTCTTAGAGATCAAAAGGGCTGTCGATCATCTATCAAGT 150
34 TyrAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50

```

THIS PAGE BLANK (USPTO)

```

151 GGCACAGATGTCACCGTGTATGGCGCCATTGGCTTGGCTTCCACCTC 200
|||||
51 GlyAlaspleuthrValMetAlaAlaIleGlyPheGlyPheLeuthrSe 67
201 GAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheAlaArgHisSerTrpSerSerValAlaPheSerLeuPheMetL 84
251 TGGCGCTGTGTGTCATGGCAATCTGCTGACGCGCTTCCGAGCCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCCTGGAGAGTGTGCATCACACATGTCATGATTCGGCTGGCC 350
|||||
101 PheProProGlyValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATTGATGCTTGTGCTGCTGATCTCATGATGCTGTCTTGGGGAAG 400
|||||
117 rThrSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyVal 134
401 TCACTTGGCGCAGTTGGTGTGATGGTGGTGGAGAGTGAACGCTTGA 450
|||||
134 alAsnLeuValGlnLeuValValMetValGluValIleThrAlaLeu 150
451 GGCACCTGAGATGTGTATCATATATCTTCAACACAGACTACACACT 500
|||||
151 GlyThrValArgMetValIleSerAsnIlePheAsnThrAspTrpHis 167
501 GAACATGATGACACTGATGCTTGGACGCTTATTTGGGCTGTCTGG 550
|||||
167 tAsnLeuMetHisIleTrpValAlaPheAlaAlaTrpPheGlyLeuSerVal 184
551 CCGGTGCTGCCAAGCCTCTACCCGAGGAGGAGGAGATTAAGATCAG 600
|||||
184 lArpCysLeuProLysProLeuProLysGlyTrpGluAspGln 200
601 ACAACAAGATACCCAGTGTGTGCTGATGCGCGCTCTCTTGTGTG 650
|||||
201 IleLalThrIleProSerIleSerAlaMetLeuGlyValAlaLeuPheL 217
651 GATGTTGTGCCAAGTTTCAACTGCTGTGCTGAGAGTTCACATGAAA 700
|||||
217 pMetPheTrpProSerPheAsnSerAlaLeuLeuArgSerProIleGlu 234
701 GGAAGATGCGGTGTACACACTGATGCTGTAGAGACAGCAGCTGTG 750
|||||
234 rGlyAsnAlaValPheAsnThrTrpTrpAlaValAlaValSerValVal 250
751 ACAAGCATCTCAGAGGTGATCTGCTGCTACCCCAAGAGAGATCACA 800
|||||
251 ThrAlaIleSerGlySerSerIleAlaHisProGlnGlyIleSerMe 267
801 GACTTATGTGCACAGCGGTGTGGCAGAGCGGTGTGGTACT 850
|||||
267 tSerTrpMetHisAsnAlaValLeuAlaGlyValAlaValAlaIlyThr 284
851 CGTGTACCGATGATCCCTTCCCGGTGGCTGCGATGAGTGGCGGTCTG 900
|||||
284 erCysHisLeuIleHisSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTTGGCTGATCTCCGTCGGGAGCAGCAATGCTGCCGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerIleGlyAlaIlyThrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATTCACACAGCTTCATGAGGCTACACTGACCT 1000
|||||
317 nArgValLeuGlyIleTrpHisSerSerValMetHisTrpAsnPheSer 334
1001 TGTGGGTGCTCTGGAGAGATCATCATGATGCTGCTGCTGCTGAT 1050
|||||
334 eubenglyLeuLeuGlyIleIleTrpIleValLeuLeuValHisHis 350

```

```

1051 ACCGTGAGCCGGCAATGATGATGGCTTCCAGCTTCCTGACCAT 1100
|||||
351 ThrValTrpAsnGlyAsnGlyMetIleGlyPheGlnValAlaLeuArgIle 367
1101 TGGGAGACTCAGCTTGGCCATGCTGATAGCTCAGCTGTGGCTTCC 1150
|||||
367 eGlyGlyPheSerLeuAlaThrThrIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGTTGCTCCCTAAATCTTAAATATGGGAAGCACCCTGACGCTAA 1200
|||||
384 hGlyLeuLeuLeuAsnLeuGlyIleTrpValArgTrpHisAlaAlaLys 400
1201 TATTTGATGACCAAGTGTGTGGAAGTTTCTCATTTGGCTGTGGAT 1250
|||||
401 TyrPheAspArgPheValAlaPheTrpLysPheProHisLeuAlaValGlu 417
1251 T 1251
417 e 417

```

seq_name: p1r2:A30405

seq_documentation_block:

erythrocyte membrane protein Rh (unknown specificity) - human

N:Alternate names: blood group Rh antigen

C:Species: Homo sapiens (man)

C:Date: 05-Oct-1990 #sequence revision 05-Oct-1990 #text change 21-Jul-2000

C:Accession: A30405; S13060; I54273; S40515; S40516; S40517; S02085; S02087; A32509;

R:Cherif-zahar, B.; Bloy, C.; Le Van Kim, C.; Blanchard, D.; Bailly, P.; Hermand, P.;

Proc. Natl. Acad. Sci. U.S.A. 87, 6243-6247, 1990

A:Title: Molecular cloning and protein structure of a human blood group Rh polypeptid

A:Reference number: A30405; MUID:90349591

A:Accession: A30405

A:Molecule type: mRNA

A:Residues: 1-417 <CHR>

A:Cross-references: GB:M44015; NID:g337398; PIDN:AAA36567.1; PID:g337399

R:Avant, N.D.; Ridgwell, K.; Tanner, M.J.A.; Anstee, D.J.

Biochem. J. 271, 821-825, 1990

A:Title: cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rh

A:Reference number: S13060; MUID:91058522

A:Accession: S13060

A:Molecule type: preliminary

A:Residues: 1-417 <RES>

A:Molecule type: mRNA

A:Cross-references: EMBL:X54534; NID:g36017; PIDN:CAA38401.1; PID:g36018

R:Kajli, E.; Umenishi, F.; Iwamoto, S.; Ikemoto, S.

Hum. Genet. 91, 157-162, 1993

A:Title: Isolation of a new cDNA clone encoding an Rh polypeptide associated with the

A:Reference number: I54273

A:Accession: I54273

A:Molecule type: translated from GB/EMBL/DBJ

A:Residues: 1-417 <RES>

A:Cross-references: GB:S57967; NID:g299049; PIDN:AA26080.1; PID:g299050

R:Mouuo, I.; Collin, Y.; Cherif-zahar, B.; Carttron, J.P.; Le van Kim, C.

Nature Genet. 5, 62-65, 1993

A:Title: Molecular genetic basis of the human Rhesus blood group system.

A:Reference number: S40515; MUID:94035121

A:Accession: S40515

A:Molecule type: preliminary

A:Status: preliminary

A:Residues: 1-301, 'F', 303-417 <MOU1>

A:Accession: S40516

A:Molecule type: mRNA

A:Residues: 1-225, 'A', 227-301, 'F', 303-417 <MOU2>

A:Accession: S40517

A:Molecule type: preliminary

A:Status: preliminary

A:Residues: 1-15, 'C', 17-59, 'I', 61-67, 'S', 69-102, 'S', 104-225, 'A', 227-301, 'F', 303-417 <

Biochem. J. 256, 1043-1046, 1988

A:Title: Protein-sequence studies on Rh-related polypeptides suggest the presence of

THIS PAGE BLANK (USPTO)

A:Reference number: S02085; MUID:89134163
 A:Accession: S02085
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-33 <AVE2>
 A>Note: this peptide was designated blood group Rh-related protein D30
 A:Accession: S02087
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-14 <AVE3>
 A>Note: this peptide was designated blood group Rh-related protein R6A32
 R:Bløy, C.; Blanchard, D.; Dahr, W.; Beyreuther, K.; Salmon, C.; Carton, J.P.
 Blood 72, 661-666, 1988
 A>Title: Determination of the N-terminal sequence of human red cell Rh(D) polypeptide an
 A:Reference number: A32509; MUID:88294325
 A:Accession: A32509
 A:Molecule type: protein
 A:Residues: 2-11,'L',13-15,'X',17 <BL0>
 R:Shoorl, A.M.; Smith, B.L.; Agre, P.
 Proc. Natl. Acad. Sci. U.S.A. 85, 4042-4045, 1988
 A>Title: Polymorphism in the M-32,000 Rh protein purified from Rh(D)-positive and -neg
 A:Reference number: A30216; MUID:88234555
 A:Accession: A30216
 A:Molecule type: protein
 A:Residues: 2-11,'X',13-21 <SAB>
 R:Suyma, K.; Goldstein, J.; Aebersold, R.; Kent, S.
 Blood 77, 411, 1991
 A>Title: Regarding the size of Rh proteins.
 A:Reference number: A61174; MUID:91091537
 A:Accession: A61174
 A:Molecule type: protein
 A:Residues: 401-407,'DI' <SU>
 C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

alignment_scores:
 Quality: 1934.00 Length: 417
 Ratio: 4.872 Gaps: 0
 Percent Similarity: 95.204 Percent Identity: 91.607

alignment_block:
 US-09-600-714-41 x A30405 ..

Align seg 1/1 to: A30405 from: 1 to: 417

```

1 ATAGAGCTTAAGTACCGCGGCTGTCTGCGGCGCTGCGCCCTGCGGCG 50
  |||||||
1 MetSerSerLysTyrProArgSerValArgArgLysLeuProLeuPhe 17
51 CTTAAAGCTGGAAGCAGCTCATCTCTCTCTCTATTTTAAACCACT 100
  |||||||
17 AleuThrLeuGluAlaAlaLeuIleLeuLeuPheTyrPheLeuThr 34
101 ATGACGCTCTTAAAGATCAAAAGGCGCTGTCGATCCATCAAGTT 150
  |||||||
34 TyrPheAlaSerLeuGluAspGlnLysGlyLeuValAlaSerTyrGln 50
151 GCGCAAGATCTGACCGTGATGGCGCCATTTGGCTGGCTCCACACTC 200
  |||||||
51 GlyIleAspLeuThrValMetAlaAlaLeuGlyLeuGlyPheLeuTh 67
201 GAGTTCCGGAGACACAGCTGAGCAGTGGCGCTTCAACCTTTCATGC 250
  |||||||
67 TAsnPheArgIleHisSerTyrPheSerValAlaPheAsnLeuPheMe 84
251 TGGCGCTTGTGTGAGTGGCAATCTGTGTGGAGCGCTTCTGACCGAG 300
  |||||||
84 euaIaLeuGlyValGlnTyrPheAlaIleLeuLeuAspGlyPheLeu 100
301 TTCCTTTGGAGAGTGTGTATCATCACTGTTCATTTGGCTGGCGCAC 350
  |||||||
101 PheProPheGlyLysValValIleThrLeuPheSerIleArgLeuAla 117
351 CATGAGTGTGTGTGGTGTGATCTCAAGTGGATCTGTCTTGGGAGAG 400
  |||||||

```

```

117 MetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134
401 TCAACTTGGCGCAGTTGGTGTGTGTCTGCTGGAGGCTGACAGCTTGA 450
  |||||||
134 AlaAsnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLe 150
451 GGCACACTGAGGATGCTCATCACTATATCTTCAACAGACTACACAT 500
  |||||||
151 GlyThrLeuArgMetValIleSerAsnIlePheAsnThrPheIleHis 167
501 GAAATGATGCACATCTACGCTTTCGACGAGCTATTTGGGCTGTGTGG 550
  |||||||
167 TAsnLeuArgIleHisPheTyrValPheAlaAlaTyrPheLeuThrVal 184
551 CCGTGTGCGCTCCAAAGCCTTACCCGAGGAGACGAGATCAAGATCAG 600
  |||||||
184 IatPheCysLeuProLysProLeuProLysGlyThrGluAspAsn 200
601 ACAGCAGATACCAAGTTTGTGTGCTGCTGCGGCGGCTTCTGTGG 650
  |||||||
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeu 217
651 GATGTTTGGCCCAAGTTTCAACTCTGCTGTGCTGAGATCCATCGAAA 700
  |||||||
217 PheThrProSerValAsnSerProLeuLeuArgSerProIleGlnA 234
701 GGAGATATGCCGTGTCAACACTATGCTGTACGACATGAGCTGTGG 750
  |||||||
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValAl 250
751 ACAGCATCTCAGGTCATCTTGGCTGCACCCCAAGGAGATCAGCA 800
  |||||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSer 267
801 GACTTATGTGCACAGTGCAGTGTGGCAGAGCGCTGCTGTGGTACCT 850
  |||||||
267 ThrTyrValHisSerAlaValLeuAlaGlyValAlaValGlyThrS 284
851 CGTGTACCTGATCCCTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTG 900
  |||||||
284 eCysHisLeuIleProSerProTyrPheAlaMetValLeuGlyLeuVal 300
901 GCTGGCTGATCTCCGTCGCGGGGAGCCAAATCACTCCGCGGTGTGTA 950
  |||||||
301 AlaGlyLeuIleSerIleGlyGlyAlaLysCysLeuProValCysCys 317
951 CCGAGTCTGGGATCCCAAGCTCCATCATGATGAGCTACAACTCACT 1000
  |||||||
317 PheValLeuGlyIleHisHisIleSerValMetHisSerIlePheSer 334
1001 TGCCTGCTCTGTGAGAGATCATCTACATTTGCTGCTGCTGCTGAT 1050
  |||||||
334 eueLeuGlyLeuLeuGlyGluIleThrTyrIleValLeuValLeuHis 350
1051 ACCGTGCGAGCGGCAATGGCATGATGGCTCCAGGCTCCCTCAGCAT 1100
  |||||||
351 ThrValTyrAsnGlnLysGlyMetIleGlyPheAlaValLeuLeuSer 367
1101 TGGGAACTCAGTGGCCATCTGATAGCTTCCACAGCTGTGCTCTGTA 1150
  |||||||
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTGCTCTAAATCTTAAATATGAGAAAGCACTCATGAGGCTAAA 1200
  |||||||
384 hGlyLeuLeuLeuAsnLeuLysIleTyrPheAlaPheValAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGCTGCTGTGANT 1250
  |||||||
401 TyrPheAspArgGlnValPheTyrPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

THIS PAGE BLANK (USPTO)

A: Molecule type: mRNA
 A: Residues: 1417 <RES>
 A: Cross-references: GB:U37052; NID:9607009; PIDN:AAA65626.1; PID:9607010

alignment_scores:
 Quality: 1908.00 Length: 417
 Ratio: 4.806 Gaps: 0
 Percent Similarity: 95.204 Percent Identity: 89.928

alignment_block:
 US-09-600-714-41 x I37075

Align seg 1/1 to: I37075 from: 1 to: 417

```

1 ATGACCTTAACTACCCGCGTCTGTCCGCGCTGCGCTGCGCTGCGG 50
|||||
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuLysAl 17
51 CCTAACCTGGAAGCAGCTCTCATTTCTCTCTCTCTCTCTCTCTCT 100
|||||
17 AleuThrLeuGluAlaLeuLeuLeuLeuPheTyrPhePheThrHis 34
|||||
101 ATGACCTTCTTAGAGATCAAAAGGCGCTGCGCATCTCTTCAAGTT 150
|||||
34 YrAspAlaSerLeuGluAspGlnGlyLeuValAlaSerTyrGlnVal 50
|||||
151 GGGCAAGATCTGACCGTGATGGCGCCATTGGCTGGGCTCTCTACCT 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThr 67
|||||
201 GAGTTCCGAGACACAGCTGGAGCGATGCGCTTCAACCTCTTCATGC 250
|||||
67 rSerPheAlaGlyHisSerTyrPheSerValAlaPheAsnLeuPheMet 84
|||||
251 TGGCGCTGGTGTGACAGTGGGCAATCCTGCTGACGCTTCTCTGAGCC 300
|||||
84 euAlaLeuGlyValGlnTrrAlaIleLeuLeuAspGlyPheLeuSerGln 100
|||||
301 TTCCTCTTGGGAAGGTGTCATCACATGTCTCAGTATTCGCTGGCCAC 350
|||||
101 PheProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
|||||
351 CATGAGTGTCTTGTGCTGCTGATCTCAGTGGATGCTGCTTGGGGAGG 400
|||||
117 rMetSerAlaLeuSerValLeuAlaSerAlaGlyAlaValLeuGlyTyrV 134
|||||
401 TCAACTGGGGGAGTGTGCTGATGCTGGTGGTGGGAGTGACAGCTTTA 450
|||||
134 AlAsnLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
|||||
451 GGCACCTGAGATGACATCAGTATATCTTCAACACAGACTACCAT 500
|||||
151 GlyThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
|||||
501 GAACATATGCAATCTACGTTGTCGACGCTATTTGGCTGTCTGGG 550
|||||
167 tAsnMetThrHisPheTyrValPheAlaAlaTyrPheGlyValThrValA 184
|||||
551 CCGTGGCTCTGGCAAGGCTTACCCGAGGAGGAGGAGGAGGAGGAGGAG 600
|||||
184 lArpPyrSLeuProLysProLeuProAspIleLysGlnAspLysAspGln 200
|||||
601 ACACCTAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTCTTG 650
|||||
201 lLeuAlaThrIleProSerLeuSerAlaMetLeuGlyThrLeuPheLeuTr 217
|||||
651 GATGTTTGCCCAAGTTTCAACTGCTCTGTCTGTGAAGTCAATCGAA 700
|||||
217 pMetPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnA 234
|||||
701 GGAGGAATCCGCTTCAACACTTACTGCTGACGAGTACGCTGGG 750
|||||

```

```

234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
|||||
751 ACAGCCATCTCAGGCGTCATCTTGGCTCACCCCCAAGGAAGATCAGCA 800
|||||
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsnMe 267
|||||
801 GACTATATGTCACAGTGGCGTGTGGCAGAGGCGCTGCTGGGTAAGCT 850
|||||
267 tThrTyrMetHisAsnAlaValLeuAlaGlyGlyValAlaValGlyThrS 284
|||||
851 CGTGCACCTGATCCCTTCTCCGCGCTGGCCATGCGTGGGCTCTGTG 900
|||||
284 eCysHisLeuIleThrSerProThrPheAlaMetValLeuGlyLeuVal 300
|||||
901 GCTGGGCTGATCTCCGTCGGGAGGAGCCAAAGTACCTGCCGCGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerIleGlyGlyAlaLysGlyLeuProGlyCysCysAs 317
|||||
951 CCGAGTGTGGGGAATCCCCACAGCTCATCATGGGCTTCAACTTCAGCT 1000
|||||
317 nArgValLeuGlyIleHisAspSerSerValMetHisTyrAsnPheSerL 334
|||||
1001 TCGTGGCTGCTGGAGAGATCATCTACATGTGCTGCTGCTGTGAT 1050
|||||
334 euLeuGlyLeuLeuGlyGlyIleThrTyrIleValLeuMetValLeuHis 350
|||||
1051 ACCGTGCGAGCCGCGCAATGCAATGATGGCTTCCAGGCTCTCTCAGCAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetValGlyPheGlnValLeuValSerThr 367
|||||
1101 TGGGGAATCAGCTTGGCCATGCTGATACCTGCACGCTGGCTGGCTGTA 1150
|||||
367 rGlyGlyLeuSerLeuAlaLeuAlaIleValAlaThrSerGlyLeuMet 384
|||||
1151 CAGGTTCTCCCTAAATCTTAAATGGAAGCAACCTCATGAGCTTAA 1200
|||||
384 hGlyIleLeuLeuAsnLeuLysIleThrLysAlaThrHisAlaAlaLys 400
|||||
1201 TAAATTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTTGGATT 1250
|||||
401 tYrPheAspArgGlnValPheThrLysPheProHisLeuAlaValGlyPh 417
|||||
1251 T 1251
417 e 417

```

seq_name: pIrf2:184434

seq_documentation_block:

Rhesus-like protein - crab-eating macaque
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
 C:accession: I84434
 R:Salvignol, I.; Blancher, A.; Calvus, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffi
 Blochem. Genet. 32, 201-221, 1994
 A:title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
 A:reference number: I37003; MUID:95085595
 A:accession: I84434
 A:status: preliminary; translated from GB/EMBL/DBJ
 A:molecule type: mRNA
 A:residues: 1-417 <RES>
 A:cross-references: GB:U37054; NID:9609515; PIDN:AAA65628.1; PID:9609516

alignment_scores:
 Quality: 1639.00 Length: 417
 Ratio: 4.313 Gaps: 0
 Percent Similarity: 91.127 Percent Identity: 78.177

alignment_block:

US-09-600-714-41 x I84434

Align seg 1/1 to: I84434 from: 1 to: 417

THIS PAGE IS ANK (USPTO)


```

1 ATGAGCTCTAAGTACCCGCGTCTGTCCGGCGTCCGCTGCCCCCTTGAGGC 50
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuTyrPAl 17
51 CCTACACTGAGACAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 100
17 aLeuThrLeuGlnAlaLeuLeuLeuLeuPhePhePhePhePhePhePheTyrT 34
101 ATGAGCGCTCTCTAGAGATCAAAAGGGCGTGGCGATCCTATCAAGTT 150
34 yRbPAlaSerLeuGlnLysPoiLysGlyLeuValAlaSerTyrGlnAla 50
151 GGCCAAATCTGACCGGTGATGGCGCCATTGGCTGGCTTCTCTACCTC 200
51 CysGlnAspLeuThrValMetAlaValLeuGlyLeuGlyPhePheThrSe 67
201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCACTCTTTCATGC 250
67 rAsnLeuArgArgAsnSerTyrSerValAlaPheAsnLeuPheLeuL 84
251 TGGCGCTTGTGTGACAGTGGCAATCCGTGAGCGGCTTCTGAGCCAG 300
84 euAlaLeuGlyValGlnTyrPAlaLeuLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTCTGGGAAGTGTGATCACAACACTGTTCAGTATTCGGCTGGCCAC 350
101 PheSerProGlyLysValValLeuLeuLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTCTTGTCCGCTGATCTCAGTGTGAGTGTCTCTTGGGGAAG 400
117 rArgSerThrThrSerMetLeuIleSerMetAsnAlaValLeuGlyLysV 134
401 TCAACTTGGCGCAGTGTGTGATGTGTGATGTGTGAGTGTGAGTGTGA 450
134 alAlaLeuAlaGlnLeuValValMetGluLeuValGlnLeuThrValPhe 150
451 GGCAACTGAGAGATGTGATCAGTATATCTTCAACACAGACTACCACT 500
151 GlyThrMetArgIleValIleTyrAsnIlePheLysIleAspTyrGlyMe 167
501 GAACATGATGCACATCTACGTTGCGCGCCTATTTGGCTGCTGTGG 550
167 tAsnMetGlnHisIleHisValPheAlaAlaTyrPheGlyLeuThrValA 184
551 CCTGTGCTCCCAAGCCTCTACCCGAGGAGAGAGAGATAAAGATCAG 600
184 lAlaTyrCysLeuProLysProLeuProLysGlyThrGlnAspLysTyrGln 200
601 ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGCGGCCCTCTCTTGTG 650
201 ThrThrThrSerProSerLeuPheAlaMetLeuGlyThrLeuPheLeuT 217
651 GATGTTCTGGCGCAGTCTCACTCTGCTGCTGCTGAGAACTCATCGAAA 700
217 pMePheThrProThrPheAsnSerAlaLeuLeuLeuAsnProIleGlnA 234
701 GGAAGATGCGGTGTCAACACTACTATGCTAGCAGTACAGCGGTG 750
234 rGlySAsnAlaValPheSerThrTyrTyrAlaLeuAlaValSerAlaVal 250
751 ACAGCATCTCAGGTCATCTTGGCTACCCCAAGGAGATCAGCAA 800
251 ThrAlaIleSerValSerSerLeuAlaHisProGlnArgLysIleAsnMe 267
801 GACTTATGTGACAGTGTGGGTGGCGAGAGGCTGGCTGGCTGACTACT 850
267 tThrTyrMetProAsnAlaGlyLeuAlaGlyValAlaValGlyAlaVal 284
851 CGTGTACCTATCCTTCTCCGTGGCTGGCCAGTGTGCTGGCTGCTGTG 900
284 eArgHisValIleHisSerProThrPAlaAlaMetValLeuGlyLeuVal 300

```

```

901 GCTGGCTGATCTCTCCGCGGAGCCCAAGTACTGCCGGGCTTTGTA 950
301 lAlaGlyLeuIleSerPheGlyAlaLysCysLeuProValCysPheAs 317
951 CCGAGTCTGGGATTTCCCAAGCTTCATCARGGGCTACACTTACGT 1000
317 nArgValLeuGlyThrIleGlnSerHisSerMetHisTyrThrPheGlyL 334
1001 TGCTGGCTGCTGTGAGAGATCATCTACATTTGCTGTGCTGTGAT 1050
334 euProAlaLeuLeuGlyGlnLeuThrTyrIleValAlaMetAlaLeuArg 350
1051 ACCGTGGACCGGCAATGCGATGATGCTTCCAGGCTCTCTCAGCAT 1100
351 ValPheThrPAlaSerSerAsnMetIleGlyPheGlnValLeuLeuSerTh 367
1101 TGGGAGCATCAGCTTGGCGCATGCGATGATGCTGCTGCTCTCTGA 1150
367 rGlyThrLeuSerLeuAlaMetAlaMetSerIleThrSerGlyLeuLeuT 384
1151 CAGGTTGCTCTTAATCTTAAATATGAAAGCACCTCAGAGGCTTAA 1200
384 hrGlyLeuLeuLeuAsnLeuLysIleTyrPylsGlyProHisValAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGCTGTGAT 1250
401 TyrPheAspAspGlnAlaPheTyrGlnPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

seq_name: p1r2:S78480

seq_documentation_block:

Rhesus blood group antigen-like protein isoform 4 - human

C/Species: Homo sapiens (man)

C/Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999

C/Accession: S78480; S26560

R/Colin, Y.

submitted to the EMBL Data Library, November 1991

A/Reference number: S78478

A/Accession: S78480

A/Molecule type: mRNA

A/Residues: 1-354 <COL>

A/Cross-References: EMBL:X63098; NID:g36019; PID:CAA4812.1; PID:g36020

R/Le Van Kim, C.; Cherif-Fahar, B.; Reynal, V.; Mouuo, I.; Lopez, M.; Carton, J.F.;

Blood 80, 1074-1078, 1992

A/Title: Multiple Rn messenger RNA isoforms are produced by alternative splicing.

A/Reference number: S26560; MUID:92360855

A/Accession: S26560

A/Molecule type: mRNA

A/Residues: 160-354 <LEV>

A/Cross-References: EMBL:X63098

A/Experimental source: tissue-type bone marrow

C/Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

alignment_scores:

Quality: 1627.00 Length: 399

Ratio: 4.771 Gaps: 3

Percent Similarity: 85.464 Percent Identity: 82.206

alignment_block:

US-09-600-714-41 x S78480

Align seg 1/1 to: S78480 from: 1 to: 354

```

1 ATGAGCTCTAAGTACCCGCGTCTGTCCGGCGTCCGCTGCCCCCTTGAGGC 50
1 MetSerSerLysTyrProArgSerValArgCysCysLeuProLeuTyrPAl 17
51 CCTACACTGAGACAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 100

```

THIS PAGE BLANK (USPTO)


```

|||||
Db 91 GAAGAGCTCTCATCTCCCTCTCTATTTTATTTTATCCCATATGACCCCTCTCTAGAGAT 150
OY 121 caaaaggcctcgtgacatcctaacaattggcaagactcgtcgtgtgagccatt 180
Db 151 CAAAAGGGGCTGTGGCATCTCTATCAAGTGCGGCAAGATCTGACCTGTAGTGGGCCCTT 210
OY 181 ggccttgagctccctcaccctcgagttcccgagacacagctgagcagctgagcctcacaac 240
Db 211 GGGTTGGGCTCTCTCACTCAAAATTTCCGAGACACAGCTGGAGCGATGGCCCTTCAAC 270
OY 241 cctctcaatcgtcgtcgtgtgtgtgcaagtggcaatccctgctggaagcgttctctgagccag 300
Db 271 CTCTTATATCTGTGGCGTTGGTGAGTGAGGCGCATCTCTGTGACGCGCTTCCGACCCAG 330
OY 301 ttcctctcggggaagggtgttcaatcacacactgttcaatctcgggtcggcccaatgagtgct 360
Db 331 TTCCCTCTCTGGGAAGGTGGTCAATCACTGTTTCAATTCGCTGGCCACCATGATGCT 390
OY 361 tltgcggtgctgatactcaatgatactgtcttctggggaaggtcaacttggcgcagctgtgtg 420
Db 391 ATGTGCGTGCTGATCTCAGAGGGGTGCTGTCTTGGGAGAGTCAACTTGGGCGAGTTGGTG 450
OY 421 gtgagtgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Db 451 GTGATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
OY 481 ttcacaagacacacacacacacacacacacacacacacacacacacacacacacacacacac 540
Db 511 TTCAACACAGACTACACACATGAACCTGAGGACCTTACGTGTGTCAGACCTATTTTGGG 570
OY 541 ctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Db 571 CTGACTGTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 630
OY 601 acagcaacatacccaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
Db 631 AGAGCAACATATACCCAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 690
OY 661 ccaagttcaactctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
Db 691 CCAAGTGTCACTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
OY 721 acctactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Db 751 ACTTACTATGCTCTTACAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
OY 781 ccccaagggaagataagcaagacttatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
Db 811 CCCCNAAGGAATCAAGCATGACTTATGTGACAGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 870
OY 841 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db 871 GTGGGTACCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 930
OY 901 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 931 GGTGGGCTGTATCTCACTGGGAGGCCAAGTGCTGCCGTGTGTGTGTGTGTGTGTGTGTGTGT 990
OY 961 gggatccccaagctcccaatcaatgagcttacaacttcaagcttgtgtgtgtgtgtgtgtgtgt 1020
Db 991 GGGATTCACACATTCCTCGCATGACCTCATCTTCATCTGCTGGGTCTGTGTGTGTGTGTGTGT 1050
OY 1021 atcaatcatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Db 1051 ATCACTTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1110
OY 1081 ttcgaagctcctcgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Db 1111 TTCCAGGTCCTCTCACTGAGCTTGGGGAACACAGCTTGGCATCTGTGTGTGTGTGTGTGTGT 1170
OY 1141 ggtctctcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200

```

```

Db 1171 GGTCTCTGACAGTTTGTCTCTAAATCTCAAAATATGGAAGACACCTCATGTGCTAA 1230
OY 1201 tatcttgatgaccaagttctcgaagttctcctcatcttgcttgatgttaa 1254
Db 1231 TATTTGATGACCAAGTTTCTGTGAAAGTTTCTCATTTGGCTGTGGATTTTAA 1284

RESULT 2
US-08-553-888A-2
; Sequence 2, Application us/08553888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1466 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-553-888A-2

Query Match 94.3%; Score 1182; DB 1; Length 1466;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

OY 1 atagagctcaagtaacccgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 60
Db 45 ATGAGCTCTTAAGTACCCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 104
OY 61 gaagagcctcaatctcctctctctatcttcttaccacatagacgtctccttaagat 120
Db 105 GAAGAGCTCTCATTTCTCTCTATTTTATACCACTATGAGCGCTTCTTAGAGAT 164
OY 121 caaaaggcctcgtgacatcctaacaagttggccaagatcgaacgtgtgtgtgtgtgtgtgtgt 180
Db 165 CAAAAGGGGCTGTGGCATCTATCAAGTGCGGCAAGATCTGACCTGTGTGTGTGTGTGTGTGT 224
OY 181 ggccttgagctccctcaccctcgagttcccgagacacagctgagcagctgtgtgtgtgtgtgtgt 240
Db 225 GGGTTGGGCTCTCTCACTCAAAATTTCCGAGACACAGCTGGAGCATGTGTGGCTTCAAC 284
OY 241 cctctcaatcgtcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
Db 285 CTCTTATATCTGTGGCGTTGGTGAGTGAGGCGCATCTCTGTGACGCGCTTCCGACCCAG 344
OY 301 ttcctctcgggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360

```

```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 345 TTCCCTCTGGAGGAGGTGATCATCACTGTTAGTATTCGGTGGCCACCATATGAGTCT 404
Qy 361 ttgtcgtgtcgtatctcaatgtagtgcgtctctgggaagtcacattggcgagttgtg 420
Db 405 ATGTCGGTGGTGAATCAGCGGGGCTGTCCTTGGGGAAGGCAACCTGGCCACTGTGGTG 464
Qy 421 gtgattgtcgtgtggaaggttacagctttaagcaacctgaagatgtagatcaatgatac 480
Db 465 GTGATGCTGCTGGTGGAGGTACAGCTTTAGGCCCTTGAAGAGTGCATATGTAATATC 524
Qy 481 ttcaacacagactacacacatgataatgacatactacgtgtctgcagacctatttggg 540
Db 525 TTCAACACAGACTACCAATGAACTGAGGCACTTCAACGTTGTCGACGCTATTTGGG 584
Qy 541 ctgtctgtggtcgtgtcgtcgtcgaagcctctaccggaaggaaggaaggaataagtcag 600
Db 585 CTGACTGTGGCTGGTGGCTGGCCAAAGCCTTACCCAAAGGAGGAGATGAATGATCAG 644
Qy 601 acagcaacagatacccaattgtctgcacatgtcggcgccctctctgtgattgtctgg 660
Db 645 AGACCAAGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTCTTGTGTGATGTTCTGG 704
Qy 661 ccaagttcaactctgtctgtcgtgagaagttcaatcgaaaggaagaaatgcgtgttcaac 720
Db 705 CCAAGTGTCACTCTCTCTGCTGAGAAGTCCAAATCCAAAGGAAGAAATGCCATGTTCAC 764
Qy 721 acctaatgtgtatgaagatgaagcgtgtgtgacacacatctcaaggtatcctgtgtcac 780
Db 765 ACCTACTATGCTCTAGAGAGTGTGTGTGACACCACTTACGAGGTATCCTTGGCTCAC 824
Qy 781 ccccaaggaagatcaagaagactatgtgacagctgcgtgtgtgcaaggaaggaagtcgt 840
Db 825 CCCCAAGGAAGATCAGCATATCTTATGTGACAGTGGGTTGGCGAGGGGCTGCT 884
Qy 841 gtgggtacactcgtgtcaactgataccctctcgtgtgtgctgcaatgtgtgtgtgtgtg 900
Db 885 GTGGGTACCTGTGTCACTGATCCCTTCTCGTGGCTTGCATGATGCTGGGTCTGTG 944
Qy 901 gctgtgtgtatctcgtcgtcgggggaagcaagtaactgcgggggttttaacagagtcgt 960
Db 945 GCTGGGTGATCTCATCTGCGGGGAGCCAAAGTCCCTGCGGTGTGTAAACCGAGTGTG 1004
Qy 961 ggaattccccacagctcatcatatggatggatcaactcaagctgtcgtgtgtgtgtgtg 1020
Db 1005 GGGATTCACCATCTCCGATATGACATCCATCTTCACTGCTGGGTCTCTTGGAGAG 1064
Qy 1021 atcaatcaactgtgtgtgtgtgtgtgtatgatacgttcggaagccggcaatgagatgtgc 1080
Db 1065 ATCACTCATATTGTGTGCTGTGCTGCTTCACTGCTGGAACGGCAATGGCATGTTGGC 1124
Qy 1081 ttcaaggtctctcctcagcatgtgggaactcaagctgtgcacgtgtgatagtctcaagtc 1140
Db 1125 TTCCAGGCTCCCTCAGATTGGGGAACCTGAGCTTGGCCATCGATGCTCTACGCTCT 1184
Qy 1141 ggtctctgaagaaggtgtgctcttaacttaataatagaagaagcaccatcagaagctaaa 1200
Db 1185 GGTCTCCGACAGGTTTGTCTTAATCTCAAAATATGGAAGCACTCATGTGGCTTAA 1244
Qy 1201 taatttgaagcaagtttcttgaaagtttccatcatgtgtgtgtgttttaa 1264
Db 1245 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGTGAATTTAA 1288

```

```

RESULT 3
US-08-476-176B-15/c
: Sequence 15, Application US/08476176B
: Patent No. 5958708
: GENERAL INFORMATION:
: APPLICANT: Hardman, No. 5958708man
: APPLICANT: Koldinger, Frank
: APPLICANT: Saldanha, Jose
: TITLE OF INVENTION: Reshaped monoclonal antibodies against an

```

```

: TITLE OF INVENTION: Immunoglobulin isotype
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5958708artis Patent Department
: STREET: 59 Route 10
: CITY: East Hanover
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07936-1080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/476,176B
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/127,721
: FILING DATE: 27-SEPTEMBER-1993
: APPLICATION NUMBER: US 07/952,802
: FILING DATE: 25-SEPTEMBER-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5958708ak, Henry P.
: REGISTRATION NUMBER: 33,200
: REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 277-5110
: TELEFAX: (908) 277-4306
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 467 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 22..447
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 79..447
: OTHER INFORMATION: /product= "heavy chain variable
: OTHER INFORMATION: region C21-Hey1"
US-08-476-176B-15

Query Match 3.1%; Score 39.4; DB 2; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.019;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 264 gcaatgggcaactcctcgtcagcgtcttccatgacagttccctctggaaggtgtcat 323
Db 405 GTATGCTGATGTTGCTGCGCGTGAAGTGCTGAACCTGGCCAGATGACGCGGGTGC 346
Qy 324 caactgttcgtatcttggtcgtcgcacacatgagatgagcttgcgtgtgtatcgaagga 383
Db 345 CTCGCTGCTCAGGCTGCTCACTCCATGATGAGCGGCTGCTGCTGGTGGGGA 286
Qy 384 tgcgtctggggaagtcacacttggcgcaattgtgtgtagtgtgtgtgtgtgtgtgtgt 440
Db 285 GGTGGCTTGGCTTGAACCTTCTGTTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

```

```

RESULT 4
US-08-127-721A-15/c
: Sequence 15, Application US/08127721A
: Patent No. 6066718
: GENERAL INFORMATION:
: APPLICANT: Hardman, No. 6066718man
: APPLICANT: Koldinger, Frank
: APPLICANT: Saldanha, Jose

```

```

1 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
2 TITLE OF INVENTION: Immunoglobulin isotype
3 NUMBER OF SEQUENCES: 55
4 CORRESPONDENCE ADDRESS:
5 ADDRESSSEE: No. 606671Bartis Patent and Trademark Department
6 STREET: 59 Route 10
7 CITY: East Hanover
8 STATE: New Jersey
9 COUNTRY: USA
10 ZIP: 07936-1080
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 OPERATING SYSTEM: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Patentin Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/127,721A
19 FILING DATE: 27-SEPTEMBER-1993
20 CLASSIFICATION: 424
21
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/952,802
24 FILING DATE: 25-SEPTEMBER-1992
25
26 ATTORNEY/AGENT INFORMATION:
27 NAME: No. 6066718ak, Henry P.
28 REGISTRATION NUMBER: 33,200
29 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (908) 277-5110
32 TELEFAX: (908) 277-4306
33
34 INFORMATION FOR SEQ ID NO: 15:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 467 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: double
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41
42 FEATURE:
43 NAME/KEY: CDS
44 LOCATION: 22..447
45
46 FEATURE:
47 NAME/KEY: mat_peptide
48 LOCATION: 79..447
49
50 OTHER INFORMATION: /product= "heavy chain variable
51 region C21-hay1"
52
53 US-08-127-721A-15
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942

```

```

TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable
OTHER INFORMATION: region C21-Hay1"
US-08-485-246A-15

Query Match          3.1%; Score 39.4; DB 3; Length 467;
Best Local Similarity 51.4%; Pred No. 0.019; 86; Indels 0; Gaps 0
Matches 91; Conservative 0; Mismatches 0;

QY      264 gcaatggagcaactctgtcgaacggttccaggaccagtccctcttggaagaagtgcatt 323
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       405 gTATGCTGTAAGTTGCTGCACGGCTGAAGTGCGCTGAACCTGGCCGACGTAGTACACGGCGGTTC 346

QY      324 caaatgttcagtatctggtcgctgccacaatgatgacctgttgttgcgtggtcgtatctcagtga 383
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       345 CTCGGTGGTCAAGCTGCTTCACCTCATGTAGGCGGGTGGGTGGTGTCCGGCGGTGA 286

QY      384 tgcctctcttggaagaagtcacaacttgagcgagcttggtgtagtgcgtgctggtgaagt 440
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       285 GGTCGCCCTTGGCGCTTGAACTTCTCGTTGTAGTTGGTGTGAAGGTGCCGGGCTGAT 229

RESULT
6
US-08-476-176B-11/c
: Sequence 11, Application us/08476176B
: Patent No. 5958708
GENERAL INFORMATION:
: APPLICANT: Hardman, No. 5958708man
: APPLICANT: Kolbinger, Frank
: APPLICANT: Saldanha, Jose
```

```

; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: Immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958708artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,176B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5958708ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..447
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 79..447
; OTHER INFORMATION: /product= "heavy chain variable
; OTHER INFORMATION: region C21-H1"
;
US-08-476-176B-11

Query Match          3.1%; Score 39.4; DB 2; Length 468;
Best Local Similarity 51.4%; Pred. No. 0.019;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 gcaatgagcaatcctctgagcagcctctcagccagctccctctggaagtgatcat 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 405 GTACTGCGAGTTGCTGCCGCTGAAGTGGCTGACCTGGCCGACAGTAGTACAGGGGGTCTC 346
QY 324 caacactgttcagttatctggcctgagccacatgagtgcttgcggtgctgatactcaatgaga 383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 345 CTCCTGTGTCAGGCTGCTCAGCTCAGCTCAGTAGGCGGCTGTGGCTGCTGCGGGGTGAA 286
QY 384 tgcgtctctggggaaggtcaacttggcgcaattggtgtgtagtgctgtgagaggt 440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 285 GGTGGCCTTGCCCTTGAACCTTCCTGTGTAGTTGTGTGTGAAGTCCCGGGCTGAT 229

RESULT 7
US-08-127-721A-11/C
; Sequence 11, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
;
```

```

; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: Immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066718artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,721A
; FILING DATE: 27-SEPTEMBER-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6066718ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22..447
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 79..447
; OTHER INFORMATION: /product= "heavy chain variable
; OTHER INFORMATION: region C21-H1"
;
US-08-127-721A-11

Query Match          3.1%; Score 39.4; DB 3; Length 468;
Best Local Similarity 51.4%; Pred. No. 0.019;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 gcaatgagcaatcctctgagcagcctctcagccagctccctctggaagtgatcat 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 405 GTACTGCGAGTTGCTGCCGCTGAAGTGGCTGACCTGGCCGACAGTAGTACAGGGGGTCTC 346
QY 324 caacactgttcagttatctggcctgagccacatgagtgcttgcggtgctgatactcaatgaga 383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 345 CTCCTGTGTCAGGCTGCTCAGCTCAGCTCAGTAGGCGGCTGTGGCTGCTGCGGGGTGAA 286
QY 384 tgcgtctctggggaaggtcaacttggcgcaattggtgtgtagtgctgtgagaggt 440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 285 GGTGGCCTTGCCCTTGAACCTTCCTGTGTAGTTGTGTGTGAAGTCCCGGGCTGAT 229

RESULT 8
US-08-485-246A-11/C
; Sequence 11, Application US/08485246A
; Patent No. 6072035
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6072035man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
;
```

```

1 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
2 TITLE OF INVENTION: Immunoglobulin Isotype
3 NUMBER OF SEQUENCES: 55
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: No. 6072035artis Patent Department
6 STREET: 59 Route 10
7 CITY: East Hanover
8 STATE: New Jersey
9 COUNTRY: USA
10 ZIP: 07936-1080
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patentin Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/485,246A
18 FILING DATE:
19 CLASSIFICATION: 424
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: 08/127,721
22 FILING DATE: 27-SEPTEMBER-1993
23 APPLICATION NUMBER: US 07/952,802
24 FILING DATE: 25-SEPTEMBER-1992
25 ATTORNEY/AGENT INFORMATION:
26 NAME: No. 6072035ak, Henry P.
27 REGISTRATION NUMBER: 33,200
28 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (908) 277-5110
31 TELEFAX: (908) 277-4306
32 INFORMATION FOR SEQ ID NO: 11:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 468 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37 TOPOLOGY: linear
38 MOLECULE TYPE: DNA (genomic)
39 FEATURE:
40 NAME/KEY: CDS
41 LOCATION: 22..447
42 FEATURE:
43 NAME/KEY: mat_peptide
44 LOCATION: 79..447
45 OTHER INFORMATION: /product= "heavy chain variable
46 OTHER INFORMATION: region C21-H1"
47 US-08-485-246A-11
48
49 Query Match 3 1%; Score 39.4; DB 3; Length 468;
50 Best Local Similarity 51.4%; Pred. No. 0.019; Mismatches 86; Indels 0; Gaps 0
51 Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0
52
53 QY 264 gcaagtggcgaacccgactgtgagcgcttccctgagccagttcccttccttgaggaaagtgtcat 323
54 | ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
55 Db 405 gtagtgcgtatgttgttgcgccctgaagtgacctgaaccctggcgcagtaattacacggcggtgmc 346
56 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
57 QY 324 caaacgttctaatcttgcgttgcgtccacaacatagtatgctttgttcggtygtcatctaatgtga 383
58 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
59 Db 345 ctccgttggtcacaggctcctacgtccatgatgagcggcgttgctgctgcgtccggcggtgaa 286
60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 QY 384 tgcctcttgaggaaagtccaacttgccgagttgtgtgatatgtgtctgtgagagt 440
62 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63 Db 285 ggtagccttgcccttgaaacttctcgttgtagttggtgtgaaagtgccggcgctgatt 229
64
65 RESULT 9
66 US-07-781-034-4
67 ; Sequence 4, Application us/07781034
68 ; Patent No. 5442050
69 ; GENERAL INFORMATION:
70 ; APPLICANT: Fishman, Jay A.
71 ; TITLE OF INVENTION: Molecular Cloning of Antlens Shared By

```

```

TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781,034
FILING DATE: 19911018
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MGH91-02A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-781-034-4

Query Match      3.0%; Score 38; DB 1; Length 1189;
Best Local Similarity 57.6%; Pred. No. 0.09;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy   393 gagggaagctcaactggcgacgttgtagtgcgtgtagtcgtgtagtgaggaagtacacctttagg 452
    |||||
Db   544 GTTGATGTGACCCAGATGAAGAGTGGCTGTGATGCGTAGTCATGTCATGTGTGT 603
    |||||

Oy   453 caactcggagatgcatcagtaatactctcaacacagactaccacatgaacatgatg 510
    | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   604 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661

RESULT 10
PCT-US92-08328-4
Sequence 4, Application PC/TUS9208328
GENERAL INFORMATION:
APPLICANT: Fishman, Jay A.
TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08328
FILING DATE: 19920930

```


US-08-332-766A-1

Sequence 1, Application US/08332766A

Patent No. 5843647

GENERAL INFORMATION:

APPLICANT: JEFFERYS, Alec J.

APPLICANT: ARMOUR, John

TITLE OF INVENTION: SIMPLE TANDEM REPEATS

NUMBER OF SEQUENCES: 125

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARY & CUSHMAN, L.L.P.

STREET: 1100 New York Avenue, N.W.

CITY: Washington

STATE: D. C.

COUNTRY: U.S.A.

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/332,766A

FILING DATE: 01-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 936052.9

FILING DATE: 21-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: BIRD, Donald J.

REGISTRATION NUMBER: 25,323

REFERENCE/DOCKET NUMBER: 217111/M94/0434/GB

TELECOMMUNICATION INFORMATION:

RESULT 12
 US-08-476-176B-17/c
 Sequence 17, Application US/08476176B
 Patent No. 5958708
 GENERAL INFORMATION:
 APPLICANT: Hardman, No. 5958708man
 APPLICANT: Kolbinger, Frank
 APPLICANT: Saldanha, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: Immunoglobulin isotype
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: No. 5958708artis Patent Department
 STREET: 59 Route 10
 CITY: East Hanover
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07936-1080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,176B
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/127,721
 FILING DATE: 27-SEPTEMBER-1993
 APPLICATION NUMBER: US 07/952,802
 FILING DATE: 25-SEPTEMBER-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5958708ak, Henry P.
 REGISTRATION NUMBER: 33,200
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 277-5110
 TELEFAX: (908) 277-4306
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:

THIS PAGE BLANK (ISPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2001, 10:49:57 : Search time 1362.11 Seconds
(without alignments)
8702.585 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctcctaagtaaccgcg.....attggtctgtgatttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Query	length	DB	ID	Description
	1	302.2	24.1		434	159	NS0044	NS9044 yv60b12.r1
	2	294.2	23.5		499	151	BF603905	BF603905 269531.MA
	3	288.2	23.0		488	181	BF602079	BF602079 267224.MA
	4	270.4	21.6		416	187	RI0548	RI0548 yf31e01.r1
	5	253.6	20.2		343	189	T84327	T84327 yd3ra05.r1
	6	243	19.4		515	145	BF191605	BF191605 239237.MA
	7	236.8	18.9		491	4	AA258801	AA258801 va67g11.r
	8	229.8	18.3		454	173	BG087038	BG087038 H3134C09
	9	208.6	16.6		491	1	AA016691	AA016691 mg90e03.r
C	10	194.6	15.5		668	117	T892255	T892255 yd3ra05.s1
C	11	188	15.0		645	117	AW553137	AW553137 l0222E01-r
	12	186.2	14.8		370	1	AA008551	AA008551 mg68a09.r
	13	183.4	14.6		400	105	AL363070	AL363070 AL363070
	14	183.2	14.6		456	150	BF557252	BF557252 UI-R-C0-g
C	15	178.2	14.2		482	22	AL632247	AL632247 tl20f10.x
C	16	173	13.8		365	149	BF510577	BF510577 UI-H-B14
C	17	167.6	13.4		492	159	NS5959	NS5959 yv60b12.s1
C	18	161.8	12.9		310	3	AA166171	AA166171 ms43d10.r
	19	157.4	12.6		434	18	AI322272	AI322272 mb64g09.y
C	20	153.6	12.2		391	187	RI0142	RI0142 yf31e01.s1
C	21	150.6	12.0		185	23	AI638737	AI638737 tl24f10.x
C	22	148.8	11.9		560	173	BG074370	BG074370 H3134C09
C	23	146	11.6		503	190	W36889	W36889 mb64g09.r1
C	24	133.2	10.6		584	156	C77974	C77974 C77974.Mous
C	25	130.8	10.4		235	1	AA031164	AA031164 ml47f09.r
C	26	126.2	10.1		539	163	BE112916	BE112916 UI-R-BJ1
C	27	125.8	10.0		430	105	AL362503	AL362503 AL362503
C	28	116	9.3		532	163	BE113605	BE113605 UI-R-BJ1
C	29	111	8.9		550	31	AV605502	AV605502 AW605502
C	30	110	8.8		624	163	BE140256	BE140256 pm2-HT0000
C	31	108	8.6		465	18	AI325413	AI325413 ms43d10.x
	32	106.8	8.5		607	106	AL564647	AL564647 AL564647
	33	103	8.2		397	4	AA222229	AA222229 mw21h12.r
	34	102.6	8.1		851	106	AL532510	AL532510 AL532510
	35	101.8	8.2		771	107	AUI21299	AUI21299 AUI21299
	36	99.8	8.0		166	1	AA033304	AA033304 ml43a02.r
	37	95.6	7.6		565	17	AI174781	AI174781 HA2497.Hu
C	38	93.6	7.5		420	18	AI324717	AI324717 mb64g09.x
C	39	92.8	7.4		440	14	AA956539	AA956539 UI-R-E1-f
C	40	91.4	7.3		438	117	AW553152	AW553152 l0222F08
C	41	87.2	7.0	1090	150	BFS80093	BFS80093 602099101	
C	42	86.8	6.9		449	189	W03969	W03969 za63f03.r1
C	43	86.2	6.9		514	24	AI789457	AI789457 uk99d12.y
C	44	85.2	6.8		532	139	BE757767	BE757767 212237.MA
C	45	85	6.8		551	145	BF155758	BF155758 PM0-HT091

ALIGNMENTS

RESULT	1
LOCUS	N59044
DEFINITION	N59044 434 bp mRNA EST 28-JAN-1997 y660b12.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:24103 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN); mRNA sequence.
ACCESSION	N59044
VERSION	N59044.1 GI:1202934
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 434)
AUTHORS	Hillier,L., Lennon,G., Becker,M., Ronaldo,M.F., Chiapelli,B., Chissee,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore

TITLE	'B., Morris,M., Parsons,M.T., Prange,C., Riffin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Treviskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Matra,M.
JOURNAL	Generation and analysis of 280,000 human expressed sequence tags
MEDLINE	Genome Res. 6 (9), 807-828 (1996)
COMMENT	97044478 Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Insert length: 780 Std Error: 0.00 Seq primer: T7 High quality sequence stop: 284.
FEATURES	Location/Qualifiers
source	1..434 /organism="Homo sapiens" /db_xref="GDB:379634.9" /db_xref="taxon:9606" /clone="IMAGE:247103" /clone_1fp="Soares fetal liver spleen JNPLS" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pUTR3D (Pharmacia) with a modified polylinker. Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAGAATTAATTAAAGACCTTTTATTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pUTR3 vector. Library went through one round of normalization. library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	78 a 114 c 117 g 113 t 12 others
ORIGIN	
Query Match	24.1%; Score 302.2; DB 159; Length 434;
Best Local Similarity	91.6%; Pred. No. 1.7e-72;
Matches 327; Conservative	0; Mismatches 29; Indels 1; Gaps 1;
OY	604 goaacagataccacagtgtgctgcacatgctggcgccctctcttcttgtagtcltcggcca 663
Db	1 GCACGATACCACCACTTTTANMACATGCTGGGCCCTCTTGCGATGTTCGGCCA 60
OY	664 agttcaactctgctcctctctagaagctcatcgaagaagaatcgacgtgttcaacc 723
Db	61 AGTGCAACTCTGCCTCTCTGGAAGTCACATCCAAAGAAAGANNCCATGTCAACAC 120
OY	724 tactatgctgtagcagtlacgltcgtgtgtagcaagcatcatalaaggtaacaccttgtaacccc 783
Db	121 TACTATGCTTAGCAGCATGATGTGGTGTGACAGCAGCATCTCAGGGTCACTCTTGCGTCACCCC 180
OY	784 caaagggaagatcagcaagaacttatgtgcacagtcggtgtgtggcaggagggcggtgtcgtg 843
Db	181 CAAAGGAAGATCACGATACCTTATGTGACAAGTGGGCTGTGGCAGGAGGCGCTGGCTTGG 240
OY	844 gtlactcgtgtcacctatccccctctccgttgcgtgttcgcatcgtgtcgtctcttggtgc- 902
Db	241 GGTAAGTGTGTCACACTATCCCTTCTCCGTGGCTGTGCATGAGTCTGGGCTTTGTGGCN 300
OY	903 ttgggtgatcttcgttcgggggagaccagtaacctcgcgggggtgtgtttaaccgagtgct 959
Db	301 TAGGCTGATCTCCATCGGGGGGANCCCAAGTGGCTGCCGATTTGGGNCCTCCAGGTCTCT 357
RESULT	2
LOCUS	BF603905 499 bp mRNA EST 13-DEC-2000
DEFINITION	2695313MARCB 3BOV Bos laurus cDNA 5', mRNA sequence.

Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
BF603905	BF603905.1	GI:11701985	EST.	Bos taurus	Smith,T.P.L., Casas,E., Stone,R.T., Heston,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Kohrer,G.A., Laegreid ,W.W. and Keefe,J.W.	Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle	Unpublished (2000)	USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov	Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
PCR Primers	FORWARD: AGGAACAGCTATGACCAT	BACKWARD: GTTTCACGTCACGACG	Plate: 50	row: D	column: 2	Seq primer: ATTTAGGTGACATATAG.	Location/Qualifiers	1..499	/organism="Bos taurus" /db_xref="taxon:9913" /clone_id="MARC_3BOV" /tissue_type="Pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; site.1: XbaI; site.2: XhoI; library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."
BASE COUNT	101 a	134 c	133 g	131 t					
ORIGIN									
Query Match	23.5%	Score 294.2;	DB 151;	Length 499;					
Best Local Similarity	74.9%	Pred. No. 2.8e-70;							
Matches 368;	Conservative 0;	Mismatches 123;	Indels 0;	Gaps 0;					
Oy	194	tcaactcgagtttcgggagacacacgtgagacagtggtgcctcaactcttcacgtctg	253						
Db	1	TCAATACCTCTTTTGGGGAGACATCTGTGAGCAGCATATGCTTCAATCTCTTCTGCTGG	60						
Oy	254	cgcgtgtgtcagtgaggcaatcctgtcctggaagcgttcctgagccagttccctctgga	313						
Db	61	TCTCTGGGGGTGACGTGGACAGCCCTGCTGAGATGCTCTCTGACCAACTCTCCCTTAA	120						
Oy	314	agggtgtcatcaactgttctagttatgttcgttcggcgacacatagtcttctgcgtgctga	373						
Db	121	AGATGCTCATCATATATGTCTCAATATTCAGAGACCCACCATATGATGCTATCATCTGTCTGA	180						
Oy	374	tctcagttgatgtcgttcttggggaaggtcaacttgcgcagttgtgtgtgatgtgtctg	433						
Db	181	TCTCGCGGGGTGTCTGCTGGGGAGAGGTCAACCTGCTCGAGCTGTGATATGAGAGCTGA	240						
Oy	434	tggaggtgacagctttaggcaacctggagagatggtcataatatacttcaacacagact	493						
Db	241	TAGAGGTGACACCCCTTCACTAGTCCACAGAGAGCTGTGAGATTAAGTACCTCGATATGACAA	300						
Oy	494	accacatgaacatgtgacacatcactacgtgttcgcagacctatttgggtgtctgtgacct	553						
Db	301	TGGATGTAAAGCATATGTATATGACACATGTGTGCGCGCTATTTTGGCGCTGACTGTGCTCT	360						
Oy	554	ggttcctgcgcaaacctctacccggaggaacggagagataaagatcagaacagacgataac	613						

DB	Accession	Length	Score	DB	Length	Score
DB	361 GCTGCTCCGGAGGCTTCCGCCACCGCATCAGAGCAAGAAAGATAGACACAGCAAGACC	420	23.0%	DB	151	488
OY	614 ccaagtgtctgtcagctgtgagcgccctctctctgtgagatgttctgccaagtccaac	673	75.7%	DB	118	1
DB	421 CCAGTTTGTTCACCACTGCGGACACCTCTCTGTGTGATATTCGGCCGAGTTTCAACT	480				
OY	674 ctgctgtctg 684					
DB	481 CTGCTCTGCTG 491					
RESULT	3					
LOCUS	BF602079	488 bp	EST	13-DEC-2000		
DEFINITION	267224 MARG 3BOV Bos taurus cDNA 5', mRNA sequence.					
ACCESSION	BF602079					
VERSION	BF602079.1	GI:11699302				
KEYWORDS	EST.					
SOURCE	cow.					
ORGANISM	Bos taurus					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;					
	Bovidae; Bovinae; Bos.					
REFERENCE	1 (bases 1 to 488)					
AUTHORS	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,					
	Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid					
	,W.W. and Keele,J.W.					
TITLE	Design and use of four pooled tissue normalized cDNA libraries for					
	Est discovery in cattle					
JOURNAL	Unpublished (2000)					
COMMENT	Contact: Smith TPL					
	USDA, ARS, US Meat Animal Research Center					
	PO Box 166, Clay Center, NE 68933-0166, USA					
	Tel: 402 762 4366					
	Fax: 402 762 4390					
	Email: smith@email.marc.usda.gov					
	Single pass sequencing. Bases called and alt_trimmed with phred					
	v0.980904.e. Vector identified by cross_match with the -mismore 18					
	and -mismatch 12 options.					
	PCR primers					
	FORWARD: AGCAACAGCATATGACCAT					
	BACKWARD: GTTTCCAGTCACGACG					
	Plate: 43 row: E column: 5					
	Seq primer: ATTAGGTGACACTATAC.					
FEATURES	Location/Qualifiers					
SOURCE	1..488					
	/organism="Bos taurus"					
	/db_xref="taxon:9913"					
	/clone_lib="MARC 3BOV"					
	/tissue_type="pooled"					
	/lab_host="DH10B"					
	/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;					
	library made from pooled tissue from marrow, alveolar					
	macrophage, ovary, fetal semitendinosus muscle, and fetal					
	longissimus muscle."					
BASE COUNT	97 a 137 c 135 g 119 t					
ORIGIN						
Query Match	23.0%; Score 288.2; DB 151; Length 488;					
Best Local Similarity	75.7%; Pred. No. 1.3e-68;					
Matches 370; Conservative 0; Mismatches 118; Indels 1; Gaps 1;						
DB	604 gcaacgataccagtttctgtcgcattgtgagcgccctctctctgtgagatgttctgcca	663				
	1 GCACGACCCCGAGTTTGTTCACCACTGCGGACACCTCTCTGTGTGATATTCGGCGG	60				
OY	664 agttcaactctgctctgtgagaagtcacatcgaaagaaatgcgtgttcaaac	723				
DB	61 AGTTCAACTCTGCTCTGCTGAGCACTTGCAAGATGAA-GAACATGCGCGGTGTCAACAC	119				
OY	724 tactatgtctgtagcagtcagcgtgtgtgacagccatcagagtcactctgtgctacccc	783				

DB	120	TACTACGCCCTTGCGGTTAGCACAGTGCACGGCATCTTAATGTCAAGCTTGAGCTACACCT	179
OY	784	caagggaagatcagaagaactatgltgcacagtcgcgtgttcgcagagagcgtgagctgtg	843
DB	180	CAAGGAGAGATCAATATGATGATCATCATCCACAGGAGAGTGTGGCAGAGAGTGTGGCTGTG	239
OY	844	ggtacctcgtgtgcacctgtatccctcttcgcgtgtgttcgcatatgtgtcgtgtgtgtgtc	903
DB	240	GGTGTCCTCTTCTTACCTGATGCTCATCTCTCTCTTGGCTTGGCATAGTGTGGATGTGTGCT	299
OY	904	ggcgcatctcgttgtggggagcgaagracctgcgcgggggtgtgtgaacggatgcttggg	963
DB	300	GGGATGATCTTCATGGGGGATTCAGATACCTGCAGAGCGGTGTCTCCACGAAAGGTGC	359
OY	964	atcccccagcagctccatcatatgtgcacaaactcgaactcgtgtgtgtgtgtgtgtgtgtg	1023
DB	360	CTCCACGACACCTACGAGGGGTGCACATCAACCTTTGGCTTCCGCGGTCTGTTGAGAGGANT	419
OY	1024	atctacattgtcgtcgt	1083
DB	420	GTCAACATCATCTCATGTATGCGCTTCAGGCCCCAGGAGATGATTAAGTCCACGCTCGGTAC	479
OY	1084	caggtccctc 1092	
DB	480	AAAGTACTC 488	
RESULT	4		
LOCUS	R10548	416 bp	mRNA EST 06-APR-1995
DEFINITION	yf1e01.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:128472 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN);, mRNA sequence.		
ACCESSION	R10548		
VERSION	R10548.1	GI:762504	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 416) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,P., Trevasakis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.		
TITLE	The Washu-Merck EST Project		
JOURNAL	Unpublished (1995)		
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 1680 High quality sequence stops: 202 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1680 Std Error: 0.00 Seq primer: M13RP1 High quality sequence stop: 202. Location/Qualifiers 1. 416 /organism="Homo sapiens" /db_xref="GDB:480633" /db_xref="taxon:9606" /clone="IMAGE:128472" /clone_lib="Soares fetal liver spleen INFLS" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pT773D (Pharmacia) with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;		
FEATURES	source		

	Query Match	21.6%	Score 270.4;	DB 187;	Length 416;	
	Best Local Similarity	97.2%;	Pred. No. 9.8e-64;			
Matches	315; Conservative	0;	Mismatches	5;	Indels	4; Gaps 4;
BASE COUNT	74 a ORIGIN	112 c 108 g 105 t	17 others			
	1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTGGCAGGATTTATTAAAGCTGTTCCTTTTTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTZ19 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."					
OY	1 atgagctctaattaccgcggcggtcgcggacgctgcgcccttgaggccaaacty 60					
Db	46 ATGACCTCTAATACCACCGGCGRCTCCNGGCCCTCGCCCCCTTGCGCCCTAACACTN 105					
OY	61 gaagcagctcatcttccttcttaatlittlaccacaatyacgcttccttagaat 120					
Db	106 GAAGCAGCTCATCTCTCTCTATTATTTTATACCACTATGACGCTTCCCTTAGAGGAT 165					
OY	121 caaaaaggggcttgtgtgaacctataagttygccaagatcbagccqytatgyggccatt 180					
Db	166 CAAAAGGGGGCTGTGGCANCTATCAAGTTGGCCAAGAATCTGACCGTAATGGCGGCATP 225					
OY	181 ggcttgggcttcccaccctcagaattccggagacacagctbgagcaftgtgccccaac 240					
Db	226 GGCTTGGGCTTCCTCTACTCCGAGTTTCCGGAGACAACAGCTGGAGCAGTGTGGCCTTYCAC 285					
OY	241 ctctcatagct-ggcgcctgtgtgtgcaqt-gggcaatcctgct-ygacgyccttcctgaqc 297					
Db	286 CTCTTCANAGCTGGCGGCTTNGETGACAGTGGGGAACATCTGCTGGAGCGGCTTCTTGAGC 345					
OY	298 cag-tttccctctyggaagtggt 320					
Db	346 CAGTTTCCCCTNTMTGGGAAGTG GT 369					
RESULT 5	T84327 LOCUS T84327 343 bp mRNA EST 16-MAR-1995					
DEFINITION	ydc37ad5.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:110384.5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN); mRNA sequence.					
VERSION	T84327 accession number T84327.1 GI:712615					
KEYWORDS	EST.					
SOURCE	Homo sapiens human.					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Hominio.					
REFERENCE AUTHORS	1 (bases 1 to 343) Hallier,L., Clark,N., Dubugue,T., Elliston,K., Hawkins,M., Holman,M., Holtzman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R. The WashU-Merck EST Project Unpublished (1995)					
JOURNAL COMMENT	Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 1034 High quality sequence stops: 295 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert length: 1034 Std Error: 0.00 Seq primer: MI3RP1 High quality sequence stop: 295.					

VERSION	T89255.1	GI:717768
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 368)	
AUTHORS	Hallier,L., Clark,N., Dubugue,T., Elliston,K., Hawkins,M., Holman	
	,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,	
	Ritkin,L., Rohlfing,T., Soares,M., Tan,F., Treaskis,E., Waterston	
	,R., Williamson,A., Wohlmann,P. and Wilson,R.	
	The Washu-Merck EST project	
TITLE	Unpublished (1995)	
JOURNAL	Contact: Wilson RK	
COMMENT	Washington University School of Medicine	
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
	Tel: 314 286 1800	
	Fax: 314 286 1810	
	Email: est@watson.wustl.edu	
	Insert Size: 1034	
	High quality sequence stops: 297 Source: IMAGE Consortium, LLNL	
	This clone is available royalty-free through LLNL ; contact the	
	IMAGE Consortium (info@image.llnl.gov) for further information.	
	Insert Length: 1034 Std Error: 0.00	
	Seq primer: -21m13	
	High quality sequence stop: 297.	
FEATURES	Location/Qualifiers	
source	1..368	/organism="Homo sapiens"
	/db_xref="GDB:466001"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:110384"	
	/clone_id="Soares fetal liver spleen INFLS"	
	/sex="male"	
	/dev_strage="20 week-post conception fetus"	
	/lab_host="DH10B (ampicillin resistant)"	
	/note="Organ: Liver and Splein; Vector: pTR73D (Pharmacia)	
	with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;	
	1st strand cDNA was primed with a Pac I - oligo(dT) primer	
	[5' AACTCGAAGATTAATTAAAGACCTTTTGTGGTTTTTTT 3'] ,	
	double-stranded cDNA was ligated to Eco RI adaptors	
	(Pharmacia), digested with Pac I and cloned into the Pac I	
	and Eco RI sites of the modified pTR73 vector. Library	
	went through one round of normalization. Library	
	constructed by Bento Soares and M.Fatima Bonaldo."	
BASE COUNT	101 a 79 c 81 g 102 t	5 others
ORIGIN		
Query Match	15.5%; Score 194.6; DB 189; Length 368;	
Best Local Similarity	93.9%; Pred. No. 7.8e-43;	
Matches 200; Conservative	0; Mismatches 13; Indels 0; Gaps 0;	
OY	1042 gtgctgataccgttcggagccggcaatggcatgatggcttccaggctcctccagcatt	1101
Db	359 GGGCMTCACTACTGTCGACGGAACGGCAATGGCATGATGGCTGCCACAGTCCCTCAGCAAT	300
OY	1102 ggggaacccaacttgccatcgtagtaactctcacaagtctgctctccagaagttgttc	1161
Db	239 GGGGACATCAGCTTGCGGCATCGTAGTACTCTCACGCTCGGGCTCTCGACAGTTGGCTC	240
OY	1162 cttaatctaaatatggaaaagcacccatcgaagctgaagcttaaatatttgatgaccagtttic	1221
Db	239 CTAAATCTCAAAATAATGTGAAGACACCCTCATGTGGCTTAATAATATTGATGATGACCAAGTTTTC	180
OY	1222 tggaaattcctcatttggtgcgtgtgatttaa	1254
Db	179 TGGAA GTTCC TCATTTGGCTGTTGGATTTTAA	147
RESULT	11	
LOCUS	AM553137.c	
	645 bp	mRNA
		EST
		31-AUG-2000

DEFINITION	L0222E01.3 Mouse Newborn Ovary cDNA Library Mus musculus cDNA clone			
ACCESSION	L0222E01.3', mRNA sequence.			
VERSION	AW553137			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (bases 1 to 645) Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Dot,H., Wood,W.H., Ilt, Becker,K.G. and Ko,M.S.H.			
TITLE	Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)			
MEDLINE	20381348			
COMMENT	Contact: George J. Kargul Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@igsun.grc.nia.nih.gov Plate: L0222 row: E column: 01 Seq primer: -21M13 Forward High quality sequence stop: 645 POLYA=Yes.			
FEATURES	location/Qualifiers			
source	1..645			
	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="taxon:10090"			
	/clone="L0222E01"			
	/clone_lib="Mouse Newborn Ovary cDNA Library"			
	/sex="female"			
	/dev_stage="Newborn Ovary"			
	/lab_host="DH10B"			
	/note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site_1: SalI; Site_2: NotI; Total RNAs were extracted from 7 Newborn Ovary. The double-stranded cDNA was synthesized by Gibco's kit with an Oligo(dt) primer [NotI primer-adapter from GibcoBRL] [5'-pGACTAGTCTACATGCAGGCGGCCGCTTTT-TTTT-3'] from 2.5ug of total RNA. The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker II-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Gentriion 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Gentriion 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Yulan Piao."			
BASE COUNT	161 a 175 c 157 g 130 t 2 others			
ORIGIN				
Query Match	15.0%: Score 188; DB 117; Length 645;			
Best Local Similarity	64.6%: Pred. No 6.2e-41;			
Matches 297; Conservative	0; Mismatches 157; Indels 6; Gaps 1;			
Qy	795 cagcaagactatgtcacacagtgcggtgttgcgcagagacgtgtgctgtggttacctcgtg 854			
Db	645 CAACATGTTGCATCATCACAATGCAGTCTGTGCGACGAGGCGCGGCGCGCCCGG 586			
Qy	855 tcaactgtaccttccttcgtgtgctgtccatagtgctgtgtgttctgtgtggtgcgtatcc 914			
Db	585 TTGGCTGATTTCTTCACCTTGATTTCCATGATGCTGTGGGCTCATATGCTGGATGATCTC 526			

DB 418 TGATCAGGCTTCTGGAGTCCGCCACTGCTGCG 455

RESULT 15

LOCUS AI632247/c

DEFINITION AI632247 482 bp mRNA EST 14-DEC-1999
 tt20f10.x1 NCI-CGAP_GC6 Homo sapiens cDNA clone IMAGE:2241355 3'
 similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN);, mRNA
 sequence.

ACCESSION AI632247
 VERSION AI632247
 KEYWORDS EST.
 SOURCE AI632247.1 GI:4683577
 ORGANISM human.

REFERENCE Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 482)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html
 Insert Length: 639 Std Error: 0.00

Seq primer: -40UP from Gibco
 High quality sequence stop: 369.

FEATURES

source

1..482

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2241355"

/clone_1ib="NCI-CGAP_GC6"

/tissue_type="Pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Plasmid DNA from the normalized library

NCI-CGAP_GC4 was prepared, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from a pool of 5,000 clones made

from the same library (clonids 1257096-1258631,

1469064-1470983, and 1475592-1476743). Subtraction by

Bento Soares and M. Fatima Bonaldo.

BASE COUNT
 ORIGIN

128 a 106 c 95 g 153 t

Query Match 14.2%; Score 178.2; DB 22; Length 482;

Best Local Similarity 98.4%; Pred. No. 2.9e-38;

Matches 180; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1072 atgattggttcacaggttcctcctcagcattgggaactcagcttgccatcgtgatagct 1131

DB 377 ACGATTGGCTTCAGGCTCCCTCCACACATTGGGGAAGTGGCCATCGATAGCT 318

OY 1132 ctacagctctggtctcctcgaaggttgctcctcctaataatgaagaagcactcat 1191

DB 317 CTACGCTCTGCTCTCGACAGGTGTGCTCAATCTCAAAATATGGAAGCACCCTCAT 258

OY 1192 gaagcctaataatttgatgacaaagtttctcgaagtccctcattgctgtgattt 1251

DB 257 GTGGCAAAATATTTCATGACCAAGTTTCTGGAAGTTTCATCATTTGGCTGTGATTT 198

OY 1252 taa 1254
 DB 197 TAA 195

Search completed: September 12, 2001, 12:18:45
 Job time: 5328 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2001, 10:49:57 ; Search time 1728.72 Seconds
(without alignments)
11220.185 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctctaagtaaccgcg.....attgtctgtgtatttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da1:*
2: gb_da2:*
3: gb_da3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_da1:*
17: gb_da2:*
18: gb_fun:*
19: gb_htgo_hum:*
20: gb_htgo_inv:*
21: gb_htgo_rod:*
22: gb_htg_hum1:*
23: gb_htg_hum2:*
24: gb_htg_hum3:*
25: gb_htg_hum4:*
26: gb_htg_hum5:*
27: gb_htg_hum6:*
28: gb_htg_hum7:*
29: gb_htg_hum8:*
30: gb_htg_inv1:*
31: gb_htg_inv2:*
32: gb_htg_other:*
33: gb_htg_rod:*
34: gb_hum1:*
35: gb_hum2:*
36: gb_hum3:*
37: gb_hum4:*
38: gb_hum5:*
39: gb_hum6:*
40: gb_hum7:*
41: gb_in:*
42: gb_om:*
43: gb_or:*

44: gb_ov:*
45: gb_pat:*
46: gb_ph:*
47: gb_pl:*
48: gb_rod:*
49: gb_sts:*
50: gb_sy:*
51: gb_un:*
52: gb_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: gb_da3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	9 AX022514	AX022514 Sequence
2	1252.4	99.9	1254	85 AB018969	AB018969 Homo sapi
3	1252.4	99.9	1354	97 HUMRHDAANTI	L08429 Human Rhd b
4	1252.4	99.9	1545	93 HSRXITI	X63094 H.sapiens m
5	1252.4	99.9	2790	93 HSRXITI	X63097 H.sapiens m
6	1250.8	99.7	1254	85 AB018966	AB018966 Homo sapi
7	1250.8	99.7	1254	85 AB018967	AB018967 Homo sapi
8	1247.6	99.5	1260	97 S78509	S78509 Rhd [human,

Db 901 GCTGGGCTGATCTCCGTCGGGGAGCCAACTACCTGCCGGGTGTGTAAACCGAGTGTCTG 960
Qy 961 gggattcccaagctccatcatatggcctacacactcagcttgcgggtctctctggagag 1020
Db 961 GGGATTCCCAAGCTCCATCATATGGCTACAACTTCAAGCTTGGGGCTCTCTGGAGAG 1020
Qy 1021 atcatctacattgtctgtctgtctgtatccgctcgagccggaatgagatattgac 1080
Db 1021 ATCATCTACATTGTGCTGTCTGTGTATACCTCGAGCCGCAATGATGATTTGGC 1080
Qy 1081 ttccaggtccctcccaagcattggggaactcagcttggccatcgatagctctcact 1140
Db 1081 TTCAGGTCCCTCCAGCATTTGGGAACCTCAGTTGGCCATCGATGATCTCAGCTCT 1140
Qy 1141 ggtctctcgaaggttgcctcctaatacttaataatgagaagcactcagagcttaa 1200
Db 1141 GGTCTCTCGAAGGTTGCTCTTAATCTTAAATATGGAAGCAACCTCATGAGGCTAAA 1200
Qy 1201 tatttgatgaccaaatttcttggaatttctcattgtctgtctgtgattaa 1254
Db 1201 TATTGTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTAA 1254

RESULT 2
AB018969
LOCUS AB018969 1254 bp mRNA PRI 03-JUL-1999
DEFINITION Homo sapiens RHDY0 mRNA for Rh blood group D antigen (RHD), complete cds.
ACCESSION AB018969.1 GI:5360898
VERSION AB018969.1
KEYWORDS RHDY0; Rh blood group D antigen (RHD).
SOURCE Homo sapiens (Isolate:YO) cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (stiles)
AUTHORS Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Taneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
TITLE Polymorphisms of RHDVa in Japanese
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 1254)
AUTHORS Uchikawa, M., Hyodo, H. and Ishikawa, Y.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1998) to the DDBJ/EMBL/Genbank databases.
Hirohbu Hyodo, Japanese Red Cross Central Blood Center, Research, 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan
(E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)

FEATURES
source
Location/Qualifiers

1..1254
/organism="Homo sapiens"
/isolate="YO"
/db_xref="taxon:9606"
1..1254
/gene="RHDY0"
1..1254
/note="RHDY0"
/note="Rh variant RHDY0"
/codon_start=1
/product="Rh blood group D antigen (RHD)"
/protein_id="BA82159.1"
/db_xref="GI:5360898"

FEATURES
translation="MSSKPRSVRYRCILPLALILEALLFFFTHYDSLDEOKL
VASYQVGDITVMAALIGLFISSFRHSSVAENFLMLALGYOMAILDGLSOP
SGKVVITLISRLATMSALSVLISVDADVILKVNADLVWVLEVTALGLRMVLSNI
FNIDYHNNMMHIVFAVFGLSVAMCLPKLPGETEDKQDTATIPSLAMGLFLFM
FMSFNSALRSPDIKRNVAENYAVAVVTAISGSLAHQGIKSYVSAVLA
GGVAVGSLRSPMLAVLGLVAGLISVGAKYLPGCCNRPVGIPIHSSIMGNSFL
IGLIGETIYVILVDIVGAGNGMIGVOVLSTIGSLALIVALTSLNLGLLNLKI
WKAPHEKATYDDVYFKFPLAVGF"
BASE COUNT 242 a 322 c 350 g 340 t
ORIGIN

Query Match 99.9%: Score 1252.4; DB 85; Length 1254;
Best Local Similarity 99.9%: Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atgagctctaagtaaccgagctctctccgagctgcgctccctctgagccataactg 60
Db 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGCGCTGCCCTCTGTGGCCCTTAACACTG 60
Qy 61 gaagcagctcattct 120
Db 61 GAAGCAGCTCATCTTCCT 120
Qy 121 caaaggggctgctgagctcctatcaagttgccaagatctgcacgggagatggccatt 180
Db 121 CAAAAGGGGCTGAGGCACTCTATCAAGTTGGCCAAAGATCTGACGGTGTATGCGGCATT 180
Qy 181 gcttgagcttcctcaccctcagcttccgagacacagcttgagcagttgagcttcaac 240
Db 181 GCTTGAGCTTCCTCACCCTCAGCTTCCGAGACACACAGCTGAGACAGTGGCTTCAAC 240
Qy 241 ctctcatgctgcgcttctgctgagtgagcaatcctctgagcggcttctctgagccag 300
Db 241 CTCTCATGCTGCGGCTGTGTGTCAGTGGCAATCTGCTGAGCGCTTCTGAGCCAG 300
Qy 301 ttccctctgggaaggtggtgctatcacactgttcaagtattcgctgcgacacatgagct 360
Db 301 TTCCCTCTGGGAAGGTGTGTATCACACTGTTCAGTATTCGGCTGCCACCATAGTGTCT 360
Qy 361 ttgcgtgctgcatcctcagtgatgagctgtctctggggaagtcacttgagcagttggt 420
Db 361 TTGCGGCTGCTGATCTCAGTGTGATGCTGTCTGGGAAGTCACTTGGCCACTTGTG 420
Qy 421 gtagtggtgctggtggaagtgaacagctttaagcaacctgagatgagtcagtaataatc 480
Db 421 GTGATGGGCTGCTGAGAGTACAGCTTTAGCAACCTGAGAGTGCATCATCAATAATC 480
Qy 481 ttcaacaagactccacatgaatgatgacatcagtcagtggttcgaagcctattttggg 540
Db 481 TTCAACAGACTACCATATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 ctgctgtgagctgtgctgtccaaagcctctaccgaggaaggaagataaagatcag 600
Db 541 CTGCTGTGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
Qy 601 acagcaacgataccaggttctgtcgaatgctggcgccctctctctgtgagatgtcttg 660
Db 601 ACAGCAACGATACCCAGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 ccaagttcaactctgctctgtcgtgagaagttcaatcgaaagaagaatgccgtgttaac 720
Db 661 CCAAGTTCAACTCTGCTGTGCTGTGAGAGTCAATCAAAAGAAAGAAAGCCGTGTTCAC 720
Qy 721 acctactatgctgtgagcagtcagtcgtgtgagcagcattcaaggtcaatctctgtcac 780
Db 721 ACCTACTATGCTGTAGCAGTCAAGCTGTGTGACACACACTCAAGGTCAATCTTGGCTCAC 780
Qy 781 ccccaagggaagatcagcaagacttatgtgacagtgcggtgtgtggaaggaagcgctgct 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGCTGTGTGCAAGAGCGGTGCT 840
Qy 841 gtgggtacctgtgttaacctatccctctctccgagctctgcaatgaggtctgtgtgtgtg 900
Db 841 GTGGGTACCTGTGTACCTGATCCCTTCTCTGAGCTGTGCAATGAGTGTGTGTGTGTG 900
Qy 901 gcttgagctgacatcctcgtctgggggaagcaagtaactgcggggtgtgttaacggagctg 960
Db 901 GCTTGAGCTGATCTCGTCTGGGGAGCAAGTACCTGCGGGGTGTGTAAACGAGTGTG 960
Qy 961 gggattcccaagctccatcatatggcctacacactcagcttgcgggtctctctggagag 1020
Db 961 GGGATTCCCAAGCTCCATCATATGGCTACAACTTCAAGCTTGGGGCTCTCTGGAGAG 1020

QY 1021 atcatctacatctgtctgtctgtctgtatgataccgctggagccggcaatggcatgtatggc 1080
|||||
Db 1021 ATCATCTACATTTGTGCTGTGTGTGATATACCGTGGAGCCGGCAATGSCATGATTTGGC 1080
QY 1081 ttccaggtctctctcagcatgtgggaactagcttggccatctgtatagcttcccaagctc 1140
|||||
Db 1081 TTCACAGTCTCTCCCTCAGCATTTGGGGAATCAGCTTGCCCTGTGATAGCTCTCACGCTCT 1140
QY 1141 ggtctccctgacaggttctgtctcccaatcttaaatatggaagaacacactcatagagcttaa 1200
|||||
Db 1141 GGCTCTCTGACAGGTTTGCCTTAATCTTAATAATGGAAGCACCCTCATGAGGCTTAA 1200
QY 1201 taatttgatgacaaagtttctggaaagttccctcaattggctgtgtgatttaa 1254
|||||
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGATTTTAA 1254

RESULT 3
HUMRHDAANTI 1354 bp mRNA PRI 24-AUG-1993
LOCUS HUMRHDAANTI 1354 bp mRNA complete cds.
DEFINITION Human Rhd blood group antigen mRNA, complete cds.
ACCESSION L08429.1 GI:337390
VERSION L08429.1 GI:337390
KEYWORDS Rh blood group; Rhd blood group; antigen; blood group antigen.
SOURCE Homo sapiens (library: HL1058b (from Clontech)) Bone marrow cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1354)
Arce,M.A., Thompson,E.S., Wagner,S., Coyne,K.E., Ferdman,B.A. and Lublin,D.M.
TITLE Molecular cloning of Rhd cDNA derived from a gene present in Rhd-positive, but not Rhd-negative individuals
JOURNAL Blood 82 (2), 651-655 (1993)
MEDLINE 93320449
FEATURES
source Location/Qualifiers
1..1354
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1254
/standard_name="Rhd"
/citation="1"
/codon_start=1
/evidence="experimental"
/product="Rhd blood group antigen"
/protein_id="AAA02679.1"
/db_xref="GI:337391"
/translation="MSSKYPBSRRCLPLMALTLERALLLFYFTHYDASLEDKCLVASVOYGDLPYMAALIGLFTLSFRHSMSVAENLEFMAIVGVMAILLDFLSQPSKVVITLFSIRLATMSALSVLISVDAYIKVNLQVLVMTLVETATLGNLMVLSINISGVDYHMNMAMHYVFAAYFGSLVAMCLPDEPDEGDKQOTITVLSLMLGALFLMFMPSNSALMRSPRIERKNVFNNTYYAVASVYTAIGESDLAPROKISKTYHSAVLAAGVAVGTSCHLIPSPMLAVGLVAGLISVGAKYLPGCCNRLVGLIPSHSINGVNSLIGLIGETIVIVLIVLDIVGAGNMIGFOVLLSIGLSLAIVALTSLGLLTGLLLMLKIWKAPHRAKTFDQYFMKPRHLYVGF"

BASE COUNT 278 a 346 c 369 g 361 t
ORIGIN

Query Match 99.9%; Score 1252.4; DB 97; Length 1354;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 121 CAAAAGGGGCTGTGSCATCTATCAAGTTGGCAGAGATCTGACCGTGTGCGGCATY 180
|||||
QY 181 ggtcttggtctctcaaccctcaggttcccgagagacaagctgagagcgtgtgcttcaac 240
|||||
Db 181 GGCTTGCGCTCTCTCACCTCAGATTTCCGAGACACAGCTGGAGCGATGTGGCTTCAAC 240
QY 241 cctctcatgctgagcttgctgtgtcagttgggcaatcctctctgagcagcttccgagccag 300
|||||
Db 241 CTCTTATCTGCGCGCTTGTGTGTGATGGGCAATCTCTCTGACCGCTTCTGACCCAG 300
QY 301 ttcccttctgggaaggttgatcatcacactgttcaaglatctcggctggccaacatagtgct 360
|||||
Db 301 TTCCTCTTGGGGAAGGTGTGATCAACACATGTTCAATTCAGATTTCGGCTGGCCACATGATGCT 360
QY 361 ttgtgtgtctgatactcaatgtagtgtgtctgttgaggaaagtgtaacttggcaggttggtg 420
|||||
Db 361 TTGTGCGCTGTATCTCAATGGATGCTGTGTGGGAAAGTCAACTTGGCGCATGTTGGTG 420
QY 421 gtagtgtgtctggtgaggtgacagctttaggcaacctgagagatggtcatcaatataatc 480
|||||
Db 421 GGTATGTGCTGTGTGAGGTGACACGCTTTAGGCAACCTGAGATGCTCATGATATATC 480
QY 481 ttcacaacagactaccacaatgaaatgatgcacatctacgtgttcgagcctatttggg 540
|||||
Db 481 TTCACACAGACTACACATGAACATGATGACATCTAGCTGTTCACACCTAATTTGGG 540
QY 541 ctgtctgtgagcctggtgctgccaagactctaacccgaggaagcaggaataaagatacag 600
|||||
Db 541 CTGTCTGTGGCTGTGTGCTGCTGCCAAGCCTTACCCGAGGAAAGAGATTAAGATTCAG 600
QY 601 acagcaacataccacagtttgtctgtccatgctggcgccctctcttctgtgatagtctgg 660
|||||
Db 601 ACAGCAACGATACCCAGTTTGTCTGTGCATGTGGGGGCCCTCTTGTGTGATGTCTGG 660
QY 661 ccaagtttcaactctgtctgtctgtgagaagttccaatcgaaagaagaatgctgttcaac 720
|||||
Db 661 CCAAGTTTCAACTCTGTCTGTGTGAGAAATCCAAATGAAAGAAAGATCCGCTGTTCAC 720
QY 721 acctactatgctgtgacagctcagcgtgtgtgacagccatctcagggatcacctctgtgac 780
|||||
Db 721 ACTTACTATGCTGTAGCATCAACGCTGTGTGACAGCATCTCAGGGTCACTTGGCTCAC 780
QY 781 ccccaagggaagatcagcaagaactatgtgcacagtgctggcttggcaggaagcgttgct 840
|||||
Db 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGCCTGTGGCAGAGGCGCT 840
QY 841 gtgggagcctgtgtacactgtatcccttccgttgctgtgacatggtgtgtgtgtgtg 900
|||||
Db 841 GTGGGTAAGCTCGTGTACCTGTGATCCCTTCCGTGGCTTGCCATGCTGGGTCTTGTG 900
QY 901 gctggctgatactcgtctggggggaagccaagtacctcgggggtgtgtgtacacgaagtctg 960
|||||
Db 901 GCTGGCTGATCTCCGCTGGGGGAGCAGCAAGTACCTCCGGGGGCTGTGTACCGAGTCTG 960
QY 961 gggatccccaagcgtccatcatgaggtgactaacctcagctgtgctgggtctgtggagag 1020
|||||
Db 961 GGATTTCCCCACAGCTCCATCATGGCTTACAACTTCCAGCTTGTGCTGTGTGGAGAG 1020
QY 1021 atcatctacatctgtctgtctgtctgtatgataccgctggagccggcaatggcatgtatggc 1080
|||||
Db 1021 ATCATCTACATTTGTGCTGTGTGTGATATACCGTGGAGCCGGCAATGSCATGATTTGGC 1080
QY 1081 ttccaggtctctcctcagcatgtgggaactagcttggccatctgtatagcttcccaagctc 1140
|||||
Db 1081 TTCACAGTCTCTCCCTCAGCATTTGGGGAATCAGCTTGCCCTGTGATAGCTCTCACGCTCT 1140
QY 1141 ggtctccctgacaggttctgtctcctaaatcttaaatatggaagaacacactcatagagcttaa 1200
|||||
Db 1141 GGCTCTCTGACAGGTTTGCCTTAATCTTAATAATGGAAGCACCCTCATGAGGCTTAA 1200
QY 1201 taatttgatgacaaagtttctggaaagttccctcaattggctgtgtgatttaa 1254
|||||

QY 1201 tatcttgatgaccagaattcttcggaagttccctcaattggtcgttgatttaa 1254
Db 1259 TATTTCATGACCAAGTTTCTTCGAACTTCATCTTGCTGCTGATTTTAA 1312

RESULT 5

HSRHXIII 2790 bp mRNA PRI 07-JUL-1997

LOCUS HSRHXIII

DEFINITION H.sapiens mRNA for rhesus polypeptide (RhXIII).

ACCESSION X63097.548953

VERSION X63097.1 GI:36046

KEYWORDS Isoform, red cell membrane protein; Rh blood group; Rh polypeptide; Rhesus polypeptide.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo. Colli, Y. Direct Submision Submitted (04-NOV-1991) Y. Colli, INSERM, U76, INTS, 6, Rue A.Cabanef, 75015 Paris, FRANCE

REFERENCE 2 (bases 1 to 2790) Le van kim, C., Cherif-Zahar, B., Raynal, V., Mouro, I., Lopez, M., Carton, J.P., and Collin, Y. Multiple Rh messenger RNA isoforms are produced by alternative splicing

TITLE Journal Blood 80 (4), 1074-1078 (1992)

JOURNAL MEDLINE 92360855

REFERENCE 3 (bases 1 to 2790) Le van kim, C., Mouro, I., Cherif-Zahar, B., Raynal, V., Cherrier, C., Carton, J.P., and Collin, Y. Molecular cloning and primary structure of the human blood group Rhd polypeptide

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)

COMMENT See also X63094-98 & M34015.

FEATURES

source Location/Qualifiers

1. 2790

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/tissue_type="bone marrow"

/clone="RhXIII"

/tissue_lib="human bone marrow cDNA"

/map="1p35"

17. 1270

/codon_start=1

/product="Rhesus polypeptide XIII"

/protein_id="CAA44811.1"

/db_xref="GI:36047"

/db_xref="SWISS-PROT:Q02161"

/translation="MSSKPPRSYRRLPLMALTLLEALLILEFFHYDASLEDOKGLVASYOYGDLTWAAIGLGLTSSFRHSSVAEFLPMALGVOMAILDDELISQFP SGKVVTLFSLRLATMSALSVLISVDAYGKVNLAQIVYVWLVETVALCNLNVISNI FNTDHYMMNMHIYVFAHYFGLSYAKCLPPLPEGTDKQDTATIPSLMGLALFLMI EWPSENSALRSPIERKNAVFNITYAAVAVTALSSSLADHPQKISTYVSAVLA GGVAAGTSCHLIPSPWLAVGLVAGLISVGAKYLPCCNRYLGIPISSKINGVNSL LGLEIRIYIVLIDVTGAGNMIGFOVLISIGLSLAIIVALTSGLLTGLLNLIK WKAPRAKFFDDQVEFKPFLAVGF"

BASE COUNT 657 a 680 c 652 g 801 t

ORIGIN

Query Match 99.9%; Score 1252.4; DB 93; Length 2790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagctctaagtaaccgggtctgtcggcgctgctccctctgggacctaaacty 60
Db 17 ATGAGCTCTAAGTAACCGCGCTGTGCCGCGCTGCTGCCCTTGGGCGCTAACACTG 76

QY 61 gaagcagctctatctctccctctatcttttaccacatagaagcttcccttaagat 120

Db 77 GAAGCAGCTCTCATTCCTCTCTATTTTATACCACATATACCTTCTTAGAGAT 136

QY 121 caaaggggtctggtgcatctatcaagttggccaagatctgacgttgaggggccatt 180

Db 137 CAAAGGGGCTGTGGATCTATCAAGTTGGCCAAAGATGACCGGATGGGGCCATT 196

QY 181 ggtctggcttccctcaccctgaatttcocgagacacagctggagcagtggtgcttcaac 240

Db 197 GCGTTGGGCTTCTTCACTGATTTCCGAGACACAGCTGGAGAGATGTGGCTTCAAC 256

QY 241 ctcttcagctggtcggttggtgtgcaagtgaggaaatccctgtgagcggctctgagccag 300

Db 257 CTCTTCATGCTGGCGCTGTGGTGCAGTGCGCAATCTGCTGCAGCGCTTCTGACCCAG 316

QY 301 ttccctctggagaggtggtgcatcacaagcttgaattatctggctggccaccatgagtct 360

Db 317 TTCCCTTCTGGAAAGGGGATCAGCTGTGAGTTTGCGTGGCCACATGAGTGTCT 376

QY 361 ttgctggtctgctcctcagtgagatgctgtctcttgagggaagttcaacttggtgcagttggt 420

Db 377 TTGTCGGTGTGATTCACAGTGAGTGTCTTGAGGAGGTCAACTTGGCGCATGAGTGTG 436

QY 421 gtgattgtctggttgagagtgacagctttaggcacactggagatggtcatcagtaataac 480

Db 437 GTGATGTGCTGTGGAGGTGACAGCTTTAGCGAACCTGAGAGTGTGATCACTAATATC 496

QY 481 ttcaacacagactaccacatgaaacatgatacagcttctgctcagcctaatttgag 540

Db 497 TTCAACACAGACTACCAATGAAACATGATGACATCTACTGCTTCCAGCTATTTTGGG 556

QY 541 ctgtctgtgctgtgtgctcgtccaaagcctctacccgagggaaacggaataaagatcag 600

Db 557 CTGTCTGTGGCTGTGTGCTGCCAAAGCCCTACCCGAGGAAACGAGATGAAGATCAG 616

QY 601 aacgcaagataccacagttgtctgcatgctggtgggcgcctctctgtggaatgcttg 660

Db 617 ACAGCAACGATACCCAGTTGTCTGCAATGCTGGGCGCTCTTCTTGTGATCTTCTGG 676

QY 661 ccaagttcaacctctctctcgtcgtgagaagctcaatcgaaagagaagatgctggttcaac 720

Db 677 CCAAGTTTCAACTCTGCTCTGCTGAGAGTCATGCAATGGAAGAAATGCGGTTCAC 736

QY 721 acctactatgctgtgacagtcagtcgtgtgagcaagcaatcgaagttcaatccttgctcac 780

Db 737 ACCTACTATGCTGTAGACAGTCAAGCTGTGTACACCATCTCAAGGTCATCTTGGCTCAC 796

QY 781 ccccaaggagaagatcagcaagacttatgtgcacaagtgaggtgtgtgcaagagcgctgct 840

Db 797 CCCCAAGGAGATCAGCAAGACTTATGTGCACAGTGGGTGTGGCAGAGGCGTGGCT 856

QY 841 gtgggtaccctcgttcaacctgaatccctctcgttggtcgttcgaatggtgctggtctgtg 900

Db 857 GTGGGTACCTCGTGTACATCGATCCCTTCCGTGGCTTCCATGATGATGCTGTTGG 916

QY 901 gctgggtcgtatctcgtcgtggggagccaaatgacctggtcgggggtgtgtgtaacgagtgctg 960

Db 917 GCTGGGCTGATCTCCGTCGGGGAGGCAAGTACTCTCCGGGGGTGTGTAAACAGAGCTG 976

QY 961 ggaattcccaagagctccacatgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1020

Db 977 GGAATTCCTCCACAGCTTCATATGGGCTTCAAGCTTGTGGGTCTGCTTGGAGAG 1036

QY 1021 atcatctacatgtgtgtcgtgtgtgtgtgatacgttcgagcgagcaatgagatgagtg 1080

Db 1037 ATCATCTACATGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1096

QY 1081 ttccaggtctctctcagcatgtgggaatcagcttggtgcatggtgataagctctcaagctct 1140

Db 1097 TTCCAGGTCTCTCTCAGCATTTGGGGAATCAGCTTGGCCATCTGTGTGTGTGTGTGTGT 1156

QY 1141 ggtctctcagcaggtgtgctcctaaatcttaaaatctgaaagcacctcaatgagtgagtg 1200

Db 1157 GGTCTCCTGACAGGTTTCCTTAATCTTAATAATATGGAAGACCTCATGAGGCTAA 1216

Qy 1201 tattttagacgaatttctggaatttctggaatttctggaatttctggaattttaa 1254
|||||

Db 1217 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCTCATTTTGGCTGTGGATTAA 1270

RESULT 6

AB018966

LOCUS AB018966 1254 bp mRNA PRI 03-JUL-1999

DEFINITION Homo sapiens Rhdva(FK) mRNA for Rh blood group D antigen (Rhd), complete cds.

ACCESSION AB018966

VERSION AB018966.1 GI:5360242

KEYWORDS Rhdva(FK); Rh blood group D antigen (RHD).

SOURCE Homo sapiens (Isolate:FK) cDNA to mRNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (sites)

AUTHORS Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Matanabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.

JOURNAL Polymorphisms of Rhdva in Japanese

REFERENCE 2 (bases 1 to 1254)

UNPUBLISHED

AUTHORS Uchikawa, M., Hyodo, H. and Ishikawa, Y.

JOURNAL Direct Submission

REFERENCE Submitted (20-OCT-1998) to the DDBJ/EMBL/GenBank databases.

REFERENCE Hironbu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cbcc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)

FEATURES

source location/Qualifiers

1..1254

/organism="Homo sapiens"

/isolate="FK"

/db_xref="taxon:9606"

1..1254

/gene="Rhdva(FK)"

1..1254

/gene="Rhdva(FK)"

/note="Rhd variant Rhdva(FK)"

/codon_start=1

/product="Rh blood group D antigen (Rhd)"

/protein_id="BA81899.1"

/db_xref="GI:5360243"

/translation="MSKRYPRSRVRCILPINALTLLEAALLLFPTTHYDASLEQKGL VASVQVQDILTVMAAIGLGLFSSFRHSMSVAFLFMALGVQMALLDGLFSOPF SGKVVITLFSIRLATMSALSVLSVDALVLSVDAVLVAVLVEVTAIGLRLVTSNI FNDYHNMNMHIYFAAYFGLSVAMCLPKPLPESTEDKQDARTPSLAMGALFLMK FMPVNSALIRSPORKNAVENTYAAVAVSVTAISGSSLAHPQKISKTYVSAVLA GYAVAGTSCILIPMLAMVIGAGLISVGAGYLPCCCRVYLGTHSSIMGTFSL LGILGELITTVLVLVDVTVGAGNGMIGFQVLISIGELSLAIVALTSLGLTLINLKI WKAPHEKRYDDVDVFMKFPHLAVGF"

BASE COUNT 241 a 323 c 351 g 339 t

ORIGIN

Query Match 99.7%; Score 1250.8; DB 85; Length 1254;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgagcttaagtaaccgcggtctctcgagcgctctgcccctctggagcctaactg 60
|||||

Db 1 ATGAGCTTAAGTACCCCGGTCTCTCGGCGCTGCTGCCCTTGAGCCCTTAACACTG 60

Qy 61 gaagcagctcattctctctctctatttlltaccacatgagcctctccttaaggat 120
|||||

Db 61 GAAGCAGCTCATCTCTCTCTATTTTACCACACTATGAGCGTCTCTTAGAGAT 120

Qy 121 caaaggggctctgagctctctataaagtggccaagatcgacgctgtagggcgcat 180
|||||

Db 121 CAAAAGGGGCTCTGTGAGCTATCAAGTTGGCCAAAGATCTGACCGTATGGCGGCATT 180

Qy 181 ggtttgggtctccctaacctcagatttccggagacacagctggagcaagtgtgaccttaac 240
|||||

Db 181 GGTTCGGGCTTCCTCCTACCTCGAGTTTCCGGAGACACACTGGAGCAGTGTGGCTTCAAC 240

Qy 241 ctcttcacgtctgagcgtctgtgtcagtgaggcaatccctctgagcagctctctcagccag 300
|||||

Db 241 CTCTTCATGCTGTGGCGCTTGTGTGAGTGGGCAATCTGTGTGAGCGCTTCTCTGAGCCAG 300

Qy 301 ttccctctgggaaggtgtgtacacactgttcagatctggctgtggccacatgagtgtc 360
|||||

Db 301 TTCCCTCTCGGGAAGGTGTGATCATCACTGTTCAGTATATTCGGCTGGCCACCATGAGTGTCT 360

Qy 361 ttgtcgtgctgagctatcagtgatgtgtcttgggggaaggtcgaactgtggcaggtgtgt 420
|||||

Db 361 TTGTGCTGTGATCTCATGTGATGCTGTCTTGGGGAAGTCACTTGGCCAGTTGGGTG 420

Qy 421 gtgatgtgctgtgaggtgtacagcctttaggcaaccgtgagatgtgtcatcagtaatac 480
|||||

Db 421 GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCAACCTGAGGATGTGATCAGTAATATC 480

Qy 481 ttcaacacagactacacacatgaaatgtacacatctcgtgttcgagcactatttggg 540
|||||

Db 481 TTCAACACAGACTACCATGATGACATGATGCAATCTACGTGTTGCGCACCTATTTTGGG 540

Qy 541 ctgtctgtgagcgtgtgtgtccaaagcctctacccggagggaacggagatgaagatcag 600
|||||

Db 541 CTGTCTGTGGCTGTGTGCTGCCAAGGCTCTACCCGAGGAGACGAGATGAAGATAG 600

Qy 601 aacgacaagataccagattgtctgtcagctgtggcgccctctcttctgtgatatctctg 660
|||||

Db 601 ACAGCAAGATACCAAGTTTGTGTCCTCAATGCTGGGCGCTCTTGTGTGATGTTCTGG 660

Qy 661 ccaagttcaactctgtctgtctgtgaaagtccaatcgaaaggaagatgctgttcaac 720
|||||

Db 661 CCAAGTTCAACTCTGTCTGTCTGTGAAAGTCCAATCGAAGGAAGAACCGGTGTTCAAC 720

Qy 721 aactactatgtctgtgagcagcagctgtgtgagcagcactcagaggtctactctgtgtcac 780
|||||

Db 721 AACTACTATGTCTGTGAGCAGCTGACGCTGTGACAGGCATCTCAGGCTCATCTTGGCTTAC 780

Qy 781 ccccaagggaagatcagaagaactatgtgtcagatgtcgatgtgttggcagagagcgctg 840
|||||

Db 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGACACTGTGGGTGTGGCAGGAGCGCTGCT 840

Qy 841 gtgggtacctgtgtcaccctatccctctctcgtgtgtgtgtgtgtgtgtgtgtgtgt 900
|||||

Db 841 GTGGGTACTCTGTGTCACTGTATCCCTTCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 900

Qy 901 gctgggctgagctcgtctgggggaagcaagtaacctgacgggggtgtgttgaacggatgtg 960
|||||

Db 901 GCTGGGCTGATCTCCGTGTGGGGAGCCAAAGTACCTGCCGGGGGTGTGTGAACCGAGTGTG 960

Qy 961 gggatctcccaagctccatcatatgggtacacactcagctgtgtgtgtgtgtgtgtgtgt 1020
|||||

Db 961 GGGATTTCCCAAGCTCCATCATATGGGTACAACTTCAGCTTGTGCTGCTGTGAGAG 1020

Qy 1021 atcatctacattgtgtctgt 1080
|||||

Db 1021 ATCATCTACATTGTGCTGT 1080

Qy 1081 ttccagctctccctcagatgtgggaactcagctgtggcatgtgtatgtatcctcagctct 1140
|||||

Db 1081 TTCCAGCTCTCCCTCAGATTTGGGAACCTGACCTTGGCCATGTGTATGCTTCACTGCTCT 1140

Qy 1141 ggtctcctgacaggtgtgtctccataatcctaataatgtgaagcaaccatcagatgagctaa 1200
|||||

Db 1141 GGTCTCTGACAGGTTTGTCTCTAATCTTAATAATATGGAAGACACTCATGAGGCTTAA 1200

Qy 1201 tattttagtgcgaagtttcttctggaagtctctcaatttggcgtgtgtgtgttttaa 1254
|||||

Db 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCTCATTTTGGCTGTGTGATTTTAA 1254

RESULT	7
LOCUS	ABO18967
DEFINITION	ABO18967 Homo sapiens RhDya(TO) mRNA for Rh blood group D antigen (RHD), complete cds.
ACCESSION	ABO18967
KEYWORDS	ABO18967.1 gi:5360244
SOURCE	RhDya(TO) : Rh blood group D antigen (RHD).
ORGANISM	Homo sapiens (Isolate:TO) cDNA to mRNA.
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.
AUTHORS	Hyodo,H., Ishikawa,Y., Kasaiyase,K., Ogawa,A., Watanabe,Y., Tsuneyama,H., Toyoda,C., Uchikawa,M., Akaza,T. and Fujii,T.
TITLE	Polymorphisms of RhDya in Japanese
JOURNAL	Unpublished (1998)
REFERENCE	2 (bases 1 to 1254)
AUTHORS	Uchikawa,M., Hyodo,H. and Ishikawa,Y.
TITLE	Direct Submission
JOURNAL	Submitted (20-OCT-1998) to the DBJ/EMBL/GenBank databases.
REFERENCE	Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research;
AUTHORS	4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan
TITLE	(E-mail:hyodo@hia.cbc.jrc.or.jp, Tel:+81-3-5485-6009,
JOURNAL	Fax:+81-3-3406-7892)
FEATURES	
source	location/Qualifiers
	1..1254
	/organism="Homo sapiens"
	/isolate="To"
	/db_xref="taxon:9606"
	1..1254
	/gene="RhDya(TO)"
	1..1254
gene	/gene="RhDya(TO)"
	/note="Rh variant RhDya(TO)"
	/codon_start=1
	/product="Rh blood group D antigen (RHD)"
	/protein_id="BAA81900.1"
	/db_xref="GI:5360245"
CDS	/translation="MSSKTPRSVRCRLPLNALTLTEALILLFYFTHYDASLEQKGLVASTOVGGDLTVMAIGLIGLTFSSFRSRHSNVAFNFMLAGVQMILIDGFLISOPFSGKVITLFLIRLATMSALSVLISVDVAIVKVNLAQVAVLVETALGNIRVVISINDENGYTHMMMHIIYVEFAVFGLSVMCKPKLPDEGTEDKDQFATPISLSAMIGALFIIMMWFPENSALLRSPLORKAMENTYYAASVVTALISSSLAHPOGKISTKYTHSVSAVLAVGGVAFNSCHLIPSPMLAMVAGIIVAGLISVGAKYLDGCCRRVIGIPHSSTMGVNFSLGLLGIEITYLYLLVDITGVANGMIGIQVLLISIELSIATLYALTSGLTRLLNLKIWKAPHEAKYFDQYFWKPHLAVE"
BASE COUNT	242 a 323 c 349 g 340 t
ORIGIN	
Query Match	99.7%; Score 1250.8; DB 85; Length 1254;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 1252:	Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy	1 atgagctcaatgaacccggcggtcgcggcgcgctgcgcgccctctgggacctaacacty 60
Db	1 ATGACTCTAACTACC CGCGGTCTTCCGGCGCTGCCTGCCCTTG GGCCCTTA CACTG 60
Oy	61 gaagcagctcctaatcttcctcttattttttaccacaatagacgcttccttagagat 120
Db	61 GAAGCAGCTCTCAATCTCTCTCTTTCTATTTTTTTACCCTACTATGACGCTTCCTTAGAGAT 120
Oy	121 caaaaggagcttgatggcatcctataagtltggccaagatctgacccgtatgtggcgcaat 180
Db	121 CAAAAGGGGCTCTGTGGCANTCTATCAAGTTGGCCAAAGATCTCACCGCTATGGCGCATTT 180
Oy	181 ggcttgggcttcctcaacctcgagtttcggagaacaacagctggagcagtggtggccttaac 240
Db	181 GGCTTGGGCTTCC TCACTCGAGTTTCCGGAGAACACAGCTGGAGCAGTGTGGCCTTTCAAC 240
Oy	241 ctatcatactgagcgcttggtgtgtagcagtgggaacatctgctgtgagcagcttcctgaagcag 300

Db	241	CTCTTCATGCTGGGGCTTGGGTGTGCATGTGGGCAATCTCTGTGACGGCTTCTCTAGCCAG	300
Qy	301	ttcccttcctgggaagctgtgtcatcaacatgttcaagtatctcgtgcctggccaacataagtgct	360
Db	301	TTCCCTTCTGTGGGAAGTGTCATCACACTGTTCAGTAATTCGGCTGGCCACCATGATGTGCT	360
Qy	361	ttgtcgtggtgcgtactcaatgagtgatgtcgtctcttggtggaaagtaaaccttggcagatgttg	420
Db	361	TTGTGGGGTGTGATCTCAGTGTAGTGTGCTGTCTTGGGAAGTCAACTGTGGCGAGTTGGTG	420
Qy	421	gttagtgtgtcgtgtggagtgtagaacgttttagcgaaccttaggatgtgtcatcaagtaatac	480
Db	421	GTGATGTGTGCGTGTGGAGGTGACAGCTTTTAGCAACCTGAGATGTGTATCAGTATATTC	480
Qy	481	ttcaacacagactaacacatgaaatgatgacatctacgtgttctgcagcctattttggg	540
Db	481	TTCAACACAGACTACCAATGAAACATGATGATGCATCATGTGTGTGCACACCTATTTTGGG	540
Qy	541	ctgtctctgtggtcgtgtgtcctggccaagccctctaccccgagggaaacggaataaagaatcag	600
Db	541	CTGTCTGTGGCTGTGGTCCGTGCCAAACCCCTCTACCCGAGGAACGGAGATTAAGATATAG	600
Qy	601	aacgacaacgataccacagtttgtctgcacatgtcgggcgcctctctcttgtagatgtctcgg	660
Db	601	ACAGCAACGATNCCCAAGTTGTCTGCCATGCTGTGGCGCCCTCTTCTGTGGATGTCTTGG	660
Qy	661	ccaagtttcaactctgtctctgtctgtagaagtccaatcgaagaagaagaatgtcgtgtctaac	720
Db	661	CCAAGTTTCAACTCTGCTGTGCTGTAGAGATGCCAATCCAAGGAAGATGCCATGTTTCAAC	720
Qy	721	aacctactatgtctgaagatcagatcgatgtgtgacagccatctcaggtgtcaatccttgtctac	780
Db	721	ACCTACTATGCTGTAGGAGATGACAGCTGTGTACAGCCATCTCAGGGATCACTTGTGGCTTAC	780
Qy	781	ccccaaaggaaagatccagcaagacttatgtgacagtgcgatgttgcaagagcgctgtgc	840
Db	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGCAGTGTGGCAGAGAGCGTGGCT	840
Qy	841	gttggttaactcgtgtcaactgtatcccttctcgtgtgtctgtgcatgtgtgtgtgttctgtg	900
Db	841	GTGGGTACTCTGTGTACACTGTATGCCCTTTCGCTGGCTTGGCATGGATGTGTGCTTGTG	900
Qy	901	gctgtggtgtatctctcgtctcggggaggaacaaagtaactctgcggggatgtgttaaccgagtcgtg	960
Db	901	GCTGTGGTGTATCTCCGTCTCGGGGGAGGCCAAATACCTCCCGGGGTGTGTAAACCGAGTGTGTG	960
Qy	961	ggagttccccaagctctcatcatgtggcgttaaaccttaagcttctgtggtctgtgttggaag	1020
Db	961	GGGATTCCCCAACACTCTCATATATGGGTATACATTTCAAGCTGTGTGGTCTGTGGACAAG	1020
Qy	1021	atcaatcaatgtgtcgtcgtgtgtgtatcaacgtctcgagccgcaatgcaatgtatgttgc	1080
Db	1021	ATCAATCATATGTGTGTGTCGTGTGATGATGACGTGTGGAGCGGCAATGGCATGATGTGGC	1080
Qy	1081	ttcgaagtccctcctcagcatgtgggaactcagcttggccatcgtgtgatctctcaagtcct	1140
Db	1081	TTCCAGGTCCTCTCACCATTTGGGAACTCAGCTTGGCATGTGTATGACTCTCACGTCCT	1140
Qy	1141	gtctctctgaaagtttgccttcttaaatcttaaatatgtgaagaagccctcatatgagcttaa	1200
Db	1141	GTCTCTCTGAAGAGTTTGCTCTTAATCTTTAAATATGTGAAGAAGCACTCATATAGGCTTAA	1200
Qy	1201	tattttgttgcacaagttctctgaaagtctccatcttgcctgttggatttttaa	1254
Db	1201	TATTTTGTATGACCAAGTTTTCTGTGAAGTTTCCATATTGGCTGTGTGGATTTTAA	1254
RESULT	8		
LOCUS	S78509	1260 bp	mrna
DEFINITION	RHD [human, D--phenotype, erythrocyte, mRNA partial, 1260 nt].		
ACCESSION	S78509		

VERSION	578509.1	GI:999309
KEYWORDS	human erythrocyte D--phenotype.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1260)	
AUTHORS	Huang,C.H., Reid,M.E. and Chen,Y.	
TITLE	Identification of a partial internal deletion in the RH locus causing the human erythrocyte D--phenotype	
JOURNAL	Blood 86 (2), 784-790 (1995)	
MEDLINE	95329738	
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gisbq 168418] from the original journal article. This sequence comes from Fig. 6.	
FEATURES	Location/Qualifiers	
source	1..1260	
gene	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	1..1254	
	/gene="Rhd"	
	1..1254	
	/gene="Rhd"	
	/note="This sequence comes from Fig. 6"	
	/codon_start=1	
	/protein_id="FAB34852.1"	
	/db_xref="GI:999310"	
CDS	/translation="MSSKYPRSVRRCPLICALEALILLFEPFTHDASLENDKGI VASQVQODLTVMALIGLGLIFSSFRHNSVSVAFNFMALGQMALLDGFLDSQFP SGKAVITLSTRLATMSALSLISVDVAVLGVNLAOLVAVLVETVALGNLTVISNIN FNDYVHNNMMLIYMFPAVFGLSVAMCPLPDETEKDDQATIPSLAMIGALFLMFA FWPFSNALRSPERKNAFENVYAAVASVTAIVSSSLAHPOGKISKTYVHSAVLVA GCVAVTSSCHLIPSPMLAVNGVLAAGISVGAAYLVGGCCRRVIGIPIHSSIMGVNFSI LGILGEIYILVLDTVGANGMIGIQVLLSIGELSLATVIALTSLTGLLLNLTNIK MKAPHVAKYEPDQYFMKFRPHAVPF"	
BASE COUNT	245 a 325 c 349 g 341 t	
ORIGIN		
Query Match	99.5%; Score 1247.6; DB 97; Length 1260;	
Best Local Similarity	99.7%; Pred. No. 0;	
Matches 1250; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
OY	1 atgagctctaaagctaccgcygctcgtccgcgcgtcctgcgccctctggccctaaacacty 60	
DB	1 ATGAGCTCTAAAGTACCGCGGTCTGCCGGCGCTGCGCCCTCTGGCCCTAACACTG 60	
OY	61 gaagcagctctcaattctctctctatttttttaccacatagacgtctccttagagat 120	
DB	61 GAAGCAGCTCTCAATCTCTCTCTTCTATTTTATTTTACCACCTATGACGCTTCTTAGAGAT 120	
OY	121 caaaagggcgctcgtgagcattcctataaatttggcgaagtcttjaccgttaatggcgccatt 180	
DB	121 CAAAAGGGCGCTGTTGGCAATCTATCAAGTTGGCCAAAGATCTGACCGTGAATGGCGCCATT 180	
OY	181 ggcctgggctctccctcaacctcgatttccggagacacagctggagcagagtggcctccaac 240	
DB	181 GGCTTGGGCTTCTCCACCTCGAGTTTCCGGAGACACACAGCTGGAGCAGTGTGGCCTTTCAAC 240	
OY	241 cctctcagcttgaggcgttggtgctgagtgaggcaatcctctgcttgagcaggtctcctgaagcag 300	
DB	241 CTTCTTCAGCTGTGGCGCTTGGTGTGCAAGTGGGCAATCTGCTGAGCGGCTTCTTGAACAG 300	
OY	301 ttcctcttgaggaaagtggctacacacattgtaatttcggcttgagccacatagagtct 360	
DB	301 TTCCCTTCTGGGAAAGTGGCTATCAACACCTGTTCAATTCGGCTTGGCCACCATGAGTGTCT 360	
OY	361 ttgctgagctatcatcagtgatgactgtctctggggaagatcaacttggcagagtgtg 420	
DB	361 TTGTGGGTCTGATCTCAAGTGAATCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420	
OY	421 gtagtgagcttgaggaggtgacagctttaagcaacttgaagatgagatgcatcagtaatac 480	
DB	421 GTATGTGAGCTTGAGGAGGTGACAGCTTTAAGCAACTTGAAGATGAGATGCAATCAATAC 480	

Db	421	GTATGAGTGGTGGTGGAGAGTATACAGCTTTTAAAGCAACCTGAGAGATGGTCAATCACTAATATC	480
Qy	481	ttcaacacagacatccacatgaaacatgatgacactctacgltgttcgagccatatttggg	540
Db	481	TTCAACACAGACTACCAATGAAATGATGACACATCTACATGTTCCGACGCTTTTWTGGG	540
Qy	541	ctgtctgtgagccttggtgtctgcacaaacccctaccggaaggaacgaggttaaagttcag	600
Db	541	CTGTCTGTGGCCTGTGGTGTCTGCCAAAGCCTCTACCCAGGAAACGGAGATTAAGATACG	600
Qy	601	acaagacagataccagattgttcctgcacatctgtgggcgccccttctctgttgaattcttg	660
Db	601	ACAGACAGATACCCAGTTTGTCTGCCATGCTGGAGCCCTCTTCTGTGATGTTCTGG	660
Qy	661	ccaagtttcaactctgtcctgtctgtagaagttccaatctgaagaagaaatgcgcgtttcac	720
Db	661	CCAAATTTCAACTCTGCTCTGTCTGAGAAGTCAATTCACAAAGGAAGATGCGTCTTCAAC	720
Qy	721	acctactatgtatgacagtcagtcgtgtgtgacagcaatctcaaggtcatccttggctcac	780
Db	721	ACCTACTATGCTGATAGACAGTGCAGGTTGGTGAACACCACATCTCAGGTCATCTTGGCTCAC	780
Qy	781	ccccagggaagatcagcaagacttatgttgcacagtcgcgtgltttgcaggagcgctggct	840
Db	781	CCCCAAGGGAAGATCAGCAAGACTTAAATGTGCACAGTCCGGGTGTGGCAGGAGCGGTGGCT	840
Qy	841	gtgggttaactcgtgtgtacacttatcccttctccgttggctttgcaatgtgtcgtgtttgtg	900
Db	841	GTGGGTACCTCTGTGTACCTTATCCCTTCTCCGTGGCTTGCATAGTGTGGGTCTTGTGTG	900
Qy	901	gctggcgtcgtatctccgtctggggagacaaagttaccctgcgcgggtgtgtttaaaccgagtctg	960
Db	901	GCTGGCGTCAATCTCCGTGGGGAGCCAAAGTACCTGCCGGGGTGTGTAAACCGAGTCTG	960
Qy	961	gggattcccccaagctccatcatatgtgtgtctaaacttaagcttgcctgggtctcgtttgagag	1020
Db	961	GGGATTCCCCAAGCTCCATCATGAGGCTACAACTTCAAGCTTGCCTGGGTCTGTTGAGAG	1020
Qy	1021	atcatctacatgttgcgtgtgtgtgtatgatacgtctcgagacgcggcaatgagcatgtatgac	1080
Db	1021	ATCATCTACATGTGTCGTCTGTCGTGCTTGATATACCGTCGAGACGGCANTGACATGATTGGC	1080
Qy	1081	ttccaggttccctccctcagcattgtgggaactcagcttggccatcgtgtatagcttcaagttc	1140
Db	1081	TTCCAGGTCTCTCCAGATTGGGGAACTCAGCTTGGCCATCGTGAATGCTTCACGCTCT	1140
Qy	1141	ggtccctcgaagcaggtgttctcctctaactttaaatatggaagaagcactcaatgagctctaa	1200
Db	1141	GGTCTCCGACAGGTTTCTCTTAATCTTAAATATAGGAAGACCTCATGTGTGCTTAA	1200
Qy	1201	tatttttgatgacaaagtttctctggaagtttccctcattgtgctgtttgatttaa	1254
Db	1201	TATTTTGAATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGGATTTTAA	1254

RESULT 9

AB018968

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AB018968

1254 bp

mRNA

Homo sapiens Rhvda(TT) mRNA for Rh blood group D antigen (RHD), complete cds.

AB018968

AB018968.1 GI:5360246

RhVda(TT); Rh blood group D antigen (RHD).

Homo sapiens (Isolate:TT) cDNA to mRNA.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (sites)

Hyodo, H., Ishikawa, Y., Kashihase, K., Ogawa, A., Matnabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.

Polymorphisms of Rhvda in Japanese

unpublished (1998)

2 (bases 1 to 1254)


```

/db_xref="taxon:9606"
/cell_type="reticulocyte"
/note="alloanti-D producing D-positive patient"
1..1254
/gene="RHD"
1..1254
/gene="RHD"
/function="associated with hemolytic disease of the
newborn in a D-positive child"
/note="This sequence comes from Fig. 4; conceptual
translation presented here differs from translation in
publication"
/codon_start=-1
/product="D category IIIC antigen"
/protein_id="AA837696.2"
/db_xref="GI:9790343"
/translaton="MSKTPRSVRCPLPLALILEALLFYEFTHYDASLEDQKL
VASVQVODLTVAALIGLFLTSSFRHSVSNVAFNLFMALGVQWAILDGLSOP
SGRVVITLFSIRLATMSAMSLISAGAVLKVLAOLVYVLEVTALGLRMVISM
FNDYHNMNMHIYFAAYFGLSYAMCLPKPLPGTEDEKDOTATIPSLAMGLFLFM
FWPSPSALLRSPILRERNNAVENVYAVAVSVTAISSSLAHOGKISKYVHSAVLA
GVAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLPCCNRVLGIPHSIMKYNLSL
LGLLGLTIVLVLVDVAGNMGIMGFVVLISIGLSLALIVALTSLGLGLINKI
MKAPHEAKYFDVDFMKFPHLAVGF"
BASE COUNT      240 a      323 c      352 g      339 t
ORIGIN
Query Match      99.4%; Score 1246; DB 97; Length 1254;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 atgagctctaagtaaccggagctgtcggcgctgcgcctccctcggccttaacctg 60
Db 1 ATGAGCTCTAAGTACCGGAGTGTGTCCGGCGTGCCTGCCCTCTGGCCCTTAACATG 60
QY 61 gaagcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 120
Db 61 GAAGCAGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 120
QY 121 caaaggagctcgtgtgacatcctcctcctcctcctcctcctcctcctcctcct 180
Db 121 CAAAGGAGCTCCTGCTGCTATCATCAAGTTCGCCAAGATCTGACCGCTGACGCT 180
QY 181 ggccttgagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 240
Db 181 GCGTGGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 240
QY 241 ccttcacatgctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
Db 241 CTTCTTCATGCTGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 301 ttcccttccttggaaggtgtgtatcatcacactgttcagltctggctggccacatgagtgct 360
Db 301 TTCCCTTCTGGGAAGGTGTGTATCATCACTGTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360
QY 361 ttgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
Db 361 ATGTGGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 421 gtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Db 421 GTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 481 ttcaacacagactacacatgaaatgtacacatcactgtgtgtgtgtgtgtgtgtgtgt 540
Db 481 TTCAACACAGACTACCAATGAAATGTACATCTACGTGTGTGTGTGTGTGTGTGTGT 540
QY 541 ctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Db 541 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 601 aacgaacagataccagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660

```

```

|||||
Db 601 ACAGCAAGATACCAGTTTGTGTGCAATGCTGGGCGCCCTTCTGTGTGATGTCTGG 660
QY 661 ccaagttcaactctgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
Db 661 CCAAGTTCAACTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 721 accacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Db 721 ACCTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 ccccaagggaagatcagcaagactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
Db 781 CCCCAAGGGAAGATCAGCAAGACTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 841 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db 841 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 901 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 961 ggaatccccaagctccatcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
Db 961 GGAATCCCCAGCTCCATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 atcatctacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Db 1021 ATCATCTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 ttccaggtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1140
Db 1081 TTCCAGGTCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1140
QY 1141 ggtcctcctgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200
Db 1141 GGTCTCCTGACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 1201 taatttgatgacaaagtttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1254
Db 1201 TATTTTGATGACAAAGTTTCTGTGAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254

RESULT 11
AF187846
LOCUS      AF187846      1458 bp      mRNA      PRI      31-OCT-1999
DEFINITION Homo sapiens RHD type IIa protein mRNA, complete cds.
ACCESSION  AF187846
VERSION     AF187846.1 GI:6164858
KEYWORDS   F1L_CDNA.
SOURCE      human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 1458)
AUTHORS   Huang,C.H., Chen,Y. and Reid,M.
TITLE     Human D(IIa) erythrocytes: Rhd protein is associated with multiple
dispersed amino acid variations
JOURNAL   Am. J. Hematol. 55 (3), 139-145 (1997)
MEDLINE   97398395
PUBMED    9256293
REFERENCE  2 (bases 1 to 1458)
AUTHORS   Huang,C.H.
TITLE     Direct Submission
JOURNAL   Submitted (18-SEP-1999) Biochemistry and Molecular Genetics, New
York Blood Center, 310 East 67th Street, New York, NY 10021, USA
FEATURES   Location/Qualifiers
source     1..1458
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="1p34-p36"

```

```

CDS
/cell_type="reticulocyte"
1..1254
/codon_start=1
/product="Rbd type IIa protein"
/db_xref="GI:6164859"
/translation="MSKYPNSVRCLPLMALTLLEAALLFFFTTHYDASLEDOKL
VASVOGDLYVMAALIGLFLTSFRHSWSSVANLMLAGVOMAILLDFLSQFP
SGKVVITLESIRLATMSALSVLSADAVGLKVNALQVVMVVEVETALGNLBMVISNI
FNTDYHMMMHIIYVFAAYEGFLVAMCLPKPEEGEDDOGATISLSAMGALFLM
FMPVNSALRSPRIERKNVFTYVAVSVYTAISGSLAPORKISRTYHSAVLA
GVAVGTSCHLIPSPMLAMVGLVAGLISVGAKTLPGCCNVNLIPHSISLNGVNSL
LGLEIETIYVLLVDYVAGNMIGFOVLLSIGELSLAIVIALISGLTLLNLKI
WRAPHEAKYFDQVFMKPEPHLAVGF"
BASE COUNT      305 a      372 c      392 g      389 t
ORIGIN
Query Match      99.1%; Score 1242.8; DB 89; Length 1458;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1247; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 atgagctctaaagtaacccggggtctgtccggcgctgtgcctcctctgggcccctaagcag 60
    |||||||
DB 1 ATGAGCTCTAAGTACCCGGCGGCTGTCCGGCGCTGCCTCCCTCGGGCTTAACACTG 60

QY 61 gaagcagctcactctcctcctctatcttttaccacataagcctcctccttagagat 120
    |||||||
DB 61 GAAGCAGCTCTACTCTCCCTCTTAATTTTAAACCACATAGCAGCTTCTTAGAGAT 120

QY 121 caaagggcgctgtgacatcctcaagttgccaagaactgacgtgatgagccact 180
    |||||||
DB 121 CAAAGGGCGCTGTGACATCTCTAAGTTGGCCAAGATCTGACCGTGATGGCGCCCTT 180

QY 181 ggtctggcctcctcactcagcttcccgagagacacagctgagcagctgtgacctcaac 240
    |||||||
DB 181 GGTCTGGCGCTCTCTCACTCTGAGTTTCCGAGACACAGCTGGAGCATGTGGCTTCAAC 240

QY 241 ctctcaatcgtgagcgttggctgtcagttggcaatcctgctgagcagctcctgaagcag 300
    |||||||
DB 241 CTCTCAATCGTGGCGCTTGGTGTGAGTGGGCAATCTGCTGAGCGGCTTCTTAGAGCAG 300

QY 301 ttcctctctgggaaggtgtgtcatcactggttcaagtattcgtgtgccaacatgaagtct 360
    |||||||
DB 301 TTCCCTCTTGGGAGGTTGGTCATCACATGTTCAATTCGCTGGCCACCATGAGTGTCT 360

QY 361 ttgtcgggtcgtatcctcagtgagtgctgctgtttgggaaagttcaacttggcgagcttgatg 420
    |||||||
DB 361 TTGTCCGCTGATCTCATGATGATGCTGTCTTGGGAAGTTCAACTTGGCGAGTTGGTG 420

QY 421 gtagatgtcgtgtagaggtgagcagctttaggcaacctgagatgggtcatcagtaaatc 480
    |||||||
DB 421 GTGATGTGCTGTGAGGTTGACAGCTTAAAGCACCTTGAGGATGTGATCACTAATATC 480

QY 481 ttcaaacagagactaacacatgatacgtatcagttgtcgcagcctattttggg 540
    |||||||
DB 481 TTCAACACAGACTAACACATGATGATGATCACTGCTGTTCCAGCCCTTATTTTGGG 540

QY 541 ctgtcgtggtcgtgtgtccgccaagcctctaccggaaggaacggaggaataagaatcaag 600
    |||||||
DB 541 CTGTCTGTGCGCTGTGTGCTGCCATGCTGGCGCCCTTCTTGTGTGATGTCTGG 600

QY 601 acagcaacgatacccaagttgtctgcatagtctggcgccctctcttctgtgagtcttg 660
    |||||||
DB 601 AGAGCAACGATACCCAGTTGTCTGCCATGCTGGCGCCCTTCTTGTGTGATGTCTGG 660

QY 661 ccaagttcaactctgtctgtcgtgagaagtccaatcgaaaggaagaatgcgcgtgttcaac 720
    |||||||
DB 661 CCAAGTGTCAACTCTGTCTGTGAGAGATCCAAATGAAAAGAAATGCGGTTCAC 720

QY 721 acctactatgcttagcagcagcgtgtagacagcatctcaggtgcatccttggtcagc 780
    |||||||
DB 721 ACCTACTATGCTTAGCAGCAGCAGTGTGTGAGACAGCCATCTCAGGGTGATCTTGCTGCTCAC 780

```

```

QY 781 ccccaagggaagatcacagcaagactatgtcacagctgctgtgttcaggaagcgtgtgct 840
    |||||||
DB 781 CCCCAAGGAAGATCACGACAAAGACTTATGTGCACAGTCCGCTGTGTGGCAGAGCGGTGCT 840

QY 841 gtgggtacctcgtgtacatgataccctctcctcgtgtgctgcaatgagtgctgtgtg 900
    |||||||
DB 841 GTGGGTACCTCGTGTACCTGATCCCTTCTCCGTGTGCTGCATGCTGTGGGTCTTGTG 900

QY 901 gctgggtatctcctcgtcgggggagcaagctactcctcgggggtgtgtgttaaccagtgctg 960
    |||||||
DB 901 GCTGGGTATCTCCCTCGTGGGGAGCAAGTACTCCTCCGGGTGTGTAAACGAGTGTG 960

QY 961 gggatccccacagcctccatgataggctacaaacttaagcttctgtggtctgttggagag 1020
    |||||||
DB 961 GGGATCCCCACAGCTCCCATGATGGGCTACAACTTACGCTGTGCTGTGGTGTGAGAG 1020

QY 1021 atcaatcaatgtgtcgtcgtgtgtgtatgataccgttcgagacggaatgcatgatatgac 1080
    |||||||
DB 1021 ATCATCTACATGTGTGCTGTGTGTGTGATACCGTCCGAGCCGCAATGAGATGTGCT 1080

QY 1081 ttccaggtctcctcactcagcatltagggaactcagcttggcactcgtgtatgactcactgct 1140
    |||||||
DB 1081 TTCCAGGTCTCTCTCAGCATTTGGGGAATCAGCTTGGCCATCTGTATGACTCTCACGCTT 1140

QY 1141 ggtctcgtacagagttgtgtcctaaacttaaaatattgaaagcacctcatgagctaa 1200
    |||||||
DB 1141 GGTCTCTGACAGAGTTGTGCTCTAATCTCAAAATATGAAACACCTCATGAGCTTAA 1200

QY 1201 tattttgagaccagatttcttggaaagttcctcattgctcgtgtgtgatttaa 1254
    |||||||
DB 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGCTGTGTGATTTTAA 1254

RESULT 12
S57971
LOCUS S57971 1251 bp mRNA PRI 28-JUN-1993
DEFINITION Rh polypeptide II (clone RnPII) [human, erythroid cells, mRNA]
ACCESSION S57971
VERSION S57971.1 GI:299051
KEYWORDS
SOURCE human erythroid cells.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Kajii F., Umenishi F., Iwamoto S. and Ikemoto S.
TITLE Isolation of a new cDNA clone encoding an Rh polypeptide associated
with the Rh blood group system
JOURNAL Hum. Genet. 91 (2), 157-162 (1993)
MEDLINE 93216282
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI glibseq 129037] from the original journal article.
This sequence comes from Fig. 3.
Map location: 1p34.3-36.1.
FEATURES
source
1..1251
Location/Qualifiers
/codon_start=1
/db_xref="taxon:9606"
1..1251
CDS
1..1251
/translation="MSKYPNSVRCLPLMALTLLEAALLFFFTTHYDASLEDOKL
VASVOGDLYVMAALIGLFLTSFRHSWSSVANLMLAGVOMAILLDFLSQFP
SGKVVITLESIRLATMSALSVLSADAVGLKVNALQVVMVVEVETALGNLBMVISNI
FNTDYHMMMHIIYVFAAYEGFLVAMCLPKPEEGEDDOGATISLSAMGALFLM
FMPVNSALRSPRIERKNVFTYVAVSVYTAISGSLAPORKISRTYHSAVLA
GVAVGTSCHLIPSPMLAMVGLVAGLISVGAKTLPGCCNVNLIPHSISLNGVNSL
LGLEIETIYVLLVDYVAGNMIGFOVLLSIGELSLAIVIALISGLTLLNLKI
WRAPHEAKYFDQVFMKPEPHLAVGF"

```

LGILGELIIVLVLDIVGANGMIGFOVLLSIGELSLAIVIALTSGLTLGLLNKI
 WKAPHEAKYFDDVFWKFPHLAVGF"
 BASE COUNT 240 a 322 c 351 g 338 t
 ORIGIN

Query Match 98.9%; Score 1239.8; DB 97; Length 1251;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 atgagctctaaagtcacccgagctgtctccgagctgtccctctggccctaaacatg 60
 Db 1 ATGAGCTCTAAGTACCCGCGGTCTCCGGCGCTGCTGCCCTCTGCGCCCTTAACATG 60
 QY 61 gaagcaactcattctcctctctatttttaaccataatgagcttctctgaagat 120
 Db 61 GAAGCACTCTATTCTCTCTCTATTCTTATTTTACCACTATGACGCTTCTCTTGAAGGAT 120
 QY 121 caaaagggctcgtgagcattcaatcgaatcgaatcgaatcgaatcgaatcgaatc 180
 Db 121 CAAAAGGGCTCGTGGGCTGCTATCAAGTTCGCCAAGATCGACCGTGAATGGCGCAT 180
 QY 181 ggttgagctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 240
 Db 181 GGTTCGCTCTCCCTCCACCTCGAGTTTCCGGAGACACAGCTGAGCACTGCTTCAAC 240
 QY 241 ctctcaatgctgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 Db 241 CTCTCAATGCTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
 QY 301 ttcctcttgaggagagtgatcactgttcaatcctcgtctgagcactgagctgt 360
 Db 301 TTCCCTCTGGGAAGGTGGTGCATCAGCTGTCTGATTTGGGCTGGCCACCATAGTGT 360
 QY 361 ttgtcgtgctgactcagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
 Db 361 TTGTCGGT 420
 QY 421 gtatggt 480
 Db 421 GTATGTTGT 480
 QY 481 ttaacaacagactaaccaatgaatgatgatgatgatgatgatgatgatgatgat 540
 Db 481 TTAACAACAGACTAACCAATGAATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 541 ctgtctgt 600
 Db 541 CTGCTGT 600
 QY 601 acaagcaagatacccaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
 Db 601 ACAGCAAGATACCCCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 QY 661 ccaagttcaactcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
 Db 661 CCAAGTTTCAACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
 QY 721 actactatctgtatgaatcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
 Db 721 ACTACTATCTGTATGAATCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 QY 781 ccccaaggaagatcagaactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
 Db 781 CCCCAGGAAGATCAGAACTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 QY 841 gtgggtacactcgt 900
 Db 841 GTGGGTACCTGCTGT 900
 QY 901 gctgggtgctgt 960
 Db 901 GCTGGGTGCTGT 960

QY 961 ggaattccccagactcattcattgagctaaactcagctgtgtgtgtgtgtgtgt 1020
 Db 961 GGAATTCCACACAGCTCATCATGAGCTTACAACTTCAAGCTTGTGTGTGTGTGTGT 1020
 QY 1021 atcatctacattgt 1080
 Db 1021 ATCATCTACATTGT 1080
 QY 1081 ttcaaggtctctcctcagcattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
 Db 1081 TTCAGGTCTCTCCACATGTGGGAAGTCAAGTTCGATCGATAGCTCTCACGCTCT 1140
 QY 1141 ggtctcctgaaggt 1200
 Db 1141 GGTCTCTGACAGGTTGT 1200
 QY 1201 tatltgatgaccaagtttctgaaatttctcattgtgtgtgtgtgtgtgtgt 1251
 Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCTATTGCTGTGTGTGTGTGT 1251

RESULT 13
 AB046420 1254 bp mRNA PRI 23-JAN-2001
 LOCUS
 DEFINITION Homo sapiens RHD mRNA for Rh blood group D antigen, complete cds.
 ACCESSION AB046420
 VERSION AB046420.1 GI:12381901
 KEYWORDS
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
 TITLE Homo sapiens RHD1 mRNA for Rh blood group antigen Rhd, complete cds
 JOURNAL Published Only in Database (2001) In press

REFERENCE
 AUTHORS Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUL-2000) Omi Toshinori, Jichi Medical School, Dept.
 of Legal Medicine and Human Genetics; 3311-1 Yakushiji
 Minamikawachi, Kawachi, Tochigi 329-0498, Japan
 (E-mail: t-omi@jichi.ac.jp, URL: www.jichi.ac.jp, Tel: 81-285-58-7342,
 Fax: 81-285-44-4902)

FEATURES
 source
 1..1254
 Location/Qualifiers

gene
 1..1254
 /db_xref="taxon:9606"
 CDS
 1..1254
 /gene="RHD"
 /gene="RHD"
 /codon_start=1
 /product="Rh blood group D antigen"
 /protein_id="BAB21261.1"
 /db_xref="GI:12381902"

/translation="MSSKYPKRSVRLPLMALTLALILFFETHYDASLEDOKGL
 VASYQVGDITVMAAIGLFTLSFRHSWSVAFLFMALGVMAILLDFLSQFP
 SKGVVITLFSIRLATMSALSVLLSVDALVKVLAOLVVVAVLEVALGRVTSNI
 FNTDYNHNMHIIIVFAAYRGLSVAMCLPRLEPGTEDEKDOTATIPISAMLGALPLFM
 FMSVNSPLRSPIDOKRNAMPNRYALAVVAVRAISGSSLAHQKRTSMRYVSAVLA
 GGVAVGTSCHLIPSPMLAVLGLVAGLISVGAKYLPGCCNRYGLPHSHIMQYNSL
 LGILGELIIVLVLDIVGANGMIGFOVLLSIGELSLAIVIALTSGLTLGLLNKI
 WKAPHEAKYFDDVFWKFPHLAVGF"
 BASE COUNT 242 a 323 c 347 g 342 t
 ORIGIN

Query Match 98.9%; Score 1239.6; DB 85; Length 1254;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1245; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY	121	caaaaggagccgfcgagacccatcaagtttggccaagaatctgaccgfcgagcgcgcat	180
Db	121	CAAAAGGAGCCGFCGAGACCCATCAAGTTTGGCCAAGATCTGACCGTAGAGGGCCCTT	180
QY	181	gacctgagcttccaccactcgaatttccggagagacacagctggagcaatgtgacctcaac	240
Db	181	GACCTTAGGCTTCCTCACCTCAAAATTTCCGGAGACAGCTGGAGCAATGTGGCTTCAC	240
QY	241	ctcttcacatgctgcgcttgtgtgtcagttgggcaatccctgtctgacggctcttcgaagcag	300
Db	241	CTCTTCACATGCTGCCTGTGGGTGTCAGATGGGCAATCTCGCTGGAGCGCTTCGAGCCAG	300
QY	301	ttccctctctgggaaggtgtgtatcaacactgttcagtatcttgctgtgccaccatgagtct	360
Db	301	TTCCCTCTCTGGGAAGGTGTGTATCACACTGTTCAGTATTCGGCTGGCCACATGAGTGTCT	360
QY	361	ttgcgcgctgcgcgactcgaatggaatgcctgtcttgggaaggccaacttggcgcaagttgtg	420
Db	361	ATGTCGCGCTGCATCTCAGCGGGTGTCTTGGGGAAAGTCAACTTGGCGCAAGTTGTG	420
QY	421	gtgatgtgtcgtgtgagagtgacagcctttagccaaccttggagatgcatcaatatac	480
Db	421	GTGATGTGTGCTGTGAGAGTGACAGCTTATGAGCACCTTGAGGATGTCATCAATATAC	480
QY	481	ttcaacacagcactcacaatgaataatgaatcaactaagtgtctgcagcctattttggg	540
Db	481	TTCAACACAGCACTCACAAATGAATAATGAATCAACTATGATGTGTGCAGCCTATTTTGGG	540
QY	541	ctgctctgtgagctgtgacctccaagccctaccacggaggaaagggagataagatcacg	600
Db	541	CTGCTCTGTGAGCTGTGACCTCCAAACCTCTACCCAGGAAAGGATTAAGATCACG	600
QY	601	acagcaacgaatacccaagtttgtctgcgaatgctggcgccctctctgtgtgattgtctg	660
Db	601	ACAGCAACGATACCAGATTTGTGTGCATGCTGGGCGCCCTCTCTGTGTGATGTTCTGG	660
QY	661	ccaagtttcaactctgtctgtctgtgagaagttccaatcgaagaagaaatgcctggttcaac	720
Db	661	CCAAGTTTCAACTCTGTCTGTCTGTAGAGTTCCAAATCCAAAGGAAGATGCCGTGTTCAAC	720
QY	721	accactatgtcttagcagatcaagctgtgtgtacacgacatccacaggtacatccttggccac	780
Db	721	ACCTACTATGTCTTAGCAGATCAAGCTGTGTGTACAGCACTTCACAGGCTCATCTTGGCTCAC	780
QY	781	ccccaaagggaagatcaagcaagaactatgtgcacagttgcagttgtctgcaggagagcgltgct	840
Db	781	CCCCAAAGGGAAGATCAAGCAAGACTTATGTGCAGCTGGGTGTGGCAGGAGCGCTGTGCT	840
QY	841	gtggagtacactgtgtacactgtatcccttccctgcgtgtgcttgcacatgtgtcgtgtgtg	900
Db	841	GTGGAGTACCTGTGTACACTGTATCCCTTCCCTGCGCTTCCATGATGCTGGGCTTGTGTG	900
QY	901	gctggagctgacatcctcgtccggggggccaagtaacctgcgggggtgtgttaacgagtgctg	960
Db	901	GCTGGAGCTGACATCCTCGTCCGGGGGACCAAGTACCTGCGGGGGTGTGTAAACCAAGTGTG	960
QY	961	ggagattcccccacagctccatcatatgtgtacaaacttcaagcttgcctgggtctcgtcttggag	1020
Db	961	GGGATTCCCCCACAGCTCCATCATGTGGGTACAACTTCAAGCTTGGGTGCTGTGGAGAG	1020
QY	1021	atcatctacaattgtcgtcgtgtgtgtatcacgctcggagccggccaatgtgcatgtatggc	1080
Db	1021	ATCATCTACAAATGTGTGTGTGTGATCACCTCGGAGCCGGCAATGTGCATGATTTGGC	1080
QY	1081	ttcagaagtcctccctcagaattgtgggaactcagcttggccatctgtagatctcgaagttc	1140
Db	1081	TTCAAGTCTCTCTCAGATTTGGGAATCTAGCTTGGCCATTCGATGAGCTCTCAGCTCT	1140
QY	1141	ggtcctcctgaaaggtttgcctctcaaatcttaaaatagagaaagcaccctcatatgagctaaa	1200
Db	1141	GGTCTCTCGAAGGTTTGTCTCTTAATCTTAAATATAGGAAGCACTCATGAGAGCTTAAA	1200
QY	1201	tatttgcatacgaagtttctctggaagtttcccatattgctgtgtgattttaa 1254	

DB	1201	TATTTTGATGACCAAGTTTCTTCGAGAGTTTCTCAATTGGCTGTGGATTTTAA	1254
RESULT	15		
LOCUS	AF037626	1247 bp	mRNA
DEFINITION	Homo sapiens truncated RHD (RHD mutant)	mRNA, complete cds.	
ACCESSION	AF037626		
VERSION	AF037626.1	GI:4104640	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.		
AUTHORS	1 (bases 1 to 1247)		
TITLE	Andrews, K.A., Walter, L.C., Saul, A. and Hyland, C.A.		
JOURNAL	The Rn D antigen negative trait in an Rh Ccee phenotypic Caucasian attributed to a four nucleotide deletion in the Rn D gene		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1247)		
TITLE	Andrews, K.A. and Hyland, C.A.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (10-DEC-1997) Malaria and Arbovirus Unit, Queensland Institute of Medical Research, 300 Herston road, Herston, QLD 4006, Australia		
FEATURES			
source	Location/Qualifiers		
	1..1247		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p34.3-p36.1"		
	1..1247		
gene	/gene="RHD"		
	/note="RHD gene with a four nucleotide deletion at the state of exon 4, resulting in a D antigen negative phenotype"		
	1..498		
CDS	/gene="RHD"		
	/codon_start=1		
	/product="truncated Rhd"		
	/protein_id="AAD02100.1"		
	/db_xref="GI:4104641"		
	/translation="MSKSPYRSVRKCLPLMALTLLEALILLFYPFTHYDASLEDQKGLVASVQVODLTVMAAIGLGLFTSFRRHSSVAFLNFMALQVMAILLDGLTSQFP SGKVIITLFSIRLATMSALSVLISYDVDAKVMALQVLYVAVLVEYVALSLRNYISNI ENTIT"		
BASE COUNT	237 a 322 c 349 g 339 t		
ORIGIN			
Query Match	98.2%; Score 1231.4; DB 88; Length 1247;		
Best Local Similarity	99.6%; Pred. No. 0; Mismatches 1; Indels 4; Gaps 1;		
Matches 1246; Conservative	0;		
1	atgagctctaaagtaaccccgagctctgcccggcgctgagcctgcctctggccctacaactg 60		
1	ATGAGCTCTAAGTACCGCGGCTGTGCTCCGGGCGTGCCTGCCTGTGGGCCCTAACA CTG 60		
61	gaagcagctctcaattctctctctctatttttttaaccaatgaagctctcttagagat 120		
61	GAAAGAGCTCTCATCTCTCTCTCTATTTTATTTTACCCACTGTGACGCTTCTTGA GAGAT 120		
121	caaaaggaggcgctgagcactctataaagttagcaaaagctggaacgtgagtgaggcgcat 180		
121	CAAAAGGGGCGCTGAGCACTCATCAATCAAGTTGGCCAAAGTGTGACCGGTGATGGCCCA TTT 180		
181	ggcttggagctctcctcaactcagattccggagacacagactgagacagttgtggcctcaac 240		
181	GGCTTGGGCTTCTCCTCACTCGAGTTTCCGGAGACACACTGTGACACAGTGTGGCTTCA AC 240		
241	ctcttcacgtctggcgcttggctgtgtgcagttgggacaatcctgtctgagcggcctctgagcag 300		
241	CTCTTCACGTCTGGCGCTTGGCTGTGTGTCAGTTGGGACAATCCTGTCTGAGCGGCCTCTG AGCAG 300		

Db 241 CTCTTCATGCTGGCGCTTGCTGTGCAGTGGCAATCCTGCTGACAGGCTTCTGAGCCAG 300
QY 301 ttcacctcttggaagtgatgcatcacactgttcaagta ttcgactgagccacatgagtgc 360
|||||
Db 301 TTCCCTTCTGGGAAGGAGGTCATCACACTGTTCAGTATTCGGCTGGCCACATGAGTGCT 360
QY 361 ttgtcggtctgatactcagtgagatgctgtcttgagggaaggtcgaacttgccgagttggtg 420
|||||
Db 361 TTGTGGGTGCTGATCTCAGTGGATGCTGTGGGGGAAGGTCAACTTGGCCAGTTGGTG 420
QY 421 gtatggtgctggttgagggtgagacagctttaggcaaccttaggagatggtcatcaatataatc 480
|||||
Db 421 GTATGAGTGTGCTGGTGGAGGTGACAGCTTTAGGCACTGAGGATGATCATCACTAATATC 480
QY 481 ttaacacagactacacacatgaaatgatactacatcagtggttcgacgcta ttttggg 540
|||||
Db 481 TTCACAC----ACTACACATGAACATGATGCATCTACGTGTTCGACGCTATTTTGGG 536
QY 541 ctgtctgtggtctggtgctgagccttcaacgaggtcaccgaggaacgagataaagatcag 600
|||||
Db 537 CTGTCTGTGGCTGTGGTGGCTGCCAAGCCCTACCCAGGGGAACGGAGATAAAGATCAG 536
QY 601 aacagcaacgatacccaagtttgctgcaatgctgagcgccctcttctgttgaatgtctgg 660
|||||
Db 597 ACAGCAACGATACCCACTTGTCTGCCATGTGGGCGCCTCTTCTGTGTGATGTCTGG 656
QY 661 ccaagtttcaactctgctcgtctgagaagttccaatcgaagaagaatgcccgtgttccaac 720
|||||
Db 657 CCAAGTTTCAACTGTGCTGTGTGAGAGATGCATGCAAGGAAGATGCGCTTTCAC 716
QY 721 acctactatgctgttagcagtcagcgtgtgtgacagccaatctcaagggtcatctccttggctaac 780
|||||
Db 717 ACCTACTATGCTGTAGCAGTCAGGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 776
QY 781 ccccaaggagaagatacgaagaagacttatgtgacagtgaggtgtgttgagagaagcggtgct 840
|||||
Db 777 CCCCAAGGGAAGATCAGCAAGACTTATGTSCAAGTGGGTGTCAGAGAGGGGCT 836
QY 841 gtgggtacacctggttcaacctgataccctctcgttggtcttgccatlggtgtggttctgtg 900
|||||
Db 837 GTGGGTACCTGCTGTGCTGATCCCTTCTCCGGGCTGCCATGGTGTGGGTCTTG 896
QY 901 gctgggtctgattctccgttcgggggagccaaatlaacctgcggggtgtgtgtaacgagtgtcg 960
|||||
Db 897 GCTGGGCTGATCTCCGTGGGGGAGCCAAATACCTGCCGGGTTGTGTAAACGAGTGTG 956
QY 961 gggattcccccaagctcccatcatalggggtatacaacttcaagcttgctggtggtcgttgagag 1020
|||||
Db 957 GGGATTCCCAAGCTCCATCATGAGGCTACAACTTCAGCTTGCTGGGTCTGCTTGGAGAG 1016
QY 1021 atcatctacatlgctgtgctgtgtgtgtatccgttcgagccggcaatggcatgaltggtc 1080
|||||
Db 1017 ATCATCTACATGTGTGCTGTGTGTGATACCGTGGAGCCGCAATGCGCATGATGTGGC 1076
QY 1081 ttcgaagttccctctcagcatgtgggaactcagcttggccaatcgtgtatagctctcaagctc 1140
|||||
Db 1077 TTCAGGTCTCTCAGCATTTGGGAACTCAGCTTGCCATCTGTGATAGCTTCACGTCT 1136
QY 1141 ggtctccctgacaggttggtcccttaaatcttaaatatggaagaacacctcalgaggtctaaa 1200
|||||
Db 1137 GGTCTCTGACAGGTTGCTCTTAATCTTAATATGAAATGAAAGCACTTCATGAGGCTAAA 1196
QY 1201 taatttgatgacaaagtttcttgaaagtttctcatttggtgttgattt 1251
|||||
Db 1197 TATTTTGATACCAAGTTTCTGGAAGTTTCTGCATTTGGCTGTGGAATTT 1247

THIS PAGE BLANK (USPTO)

```
; Sequence 41, Application US/09600714
; GENERAL INFORMATION:
; APPLICANT: Flegel, Willy A.
; APPLICANT: DRK Blutspendedienst Baden-Wuerttemberg gGmbH
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES CORRELATED
; FILE REFERENCE: 12086-002001
; CURRENT APPLICATION NUMBER: US/09/600,714
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/EP98/08319
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Homosapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1251)
09-600714-1
MSKRYRSVRRCPLPMAILIEAAILLFFFTTHYDASLEDOKGLVASYOGDLYTMAAIGLGFITSSR
RHSWSSVAENLEMLALGVOMAILLDGFLSQFPGKAVITLFSIRLATMSALSVLISVDVGLGVNLAQLV
VMVLVEVTALGNLRMYISNIFNTDYMNMNMHIYFAAYFGLVAMCLPKPLPEGTEDKDOTATIPSLSAM
LGLFLFMMFPSPNSALLRSPIERKNNAVNTYAAVSVYTAISGSSLAHPQGIKSTYGHSAVLPGVA
VGTSCHLIPSPMLAMVIGIAGLISVGAAYLPGCCNRVYIGIPHSSIMGYNFSLGLLEIITVILVID
TVGAGNGMIGFOVLISIGELSLAIVIALTSGLLTGILLNKIKAPHEAKYFPDDQYFWKFPHLAVGFXI
```

```
; Sequence 41, Application US/09600714
; GENERAL INFORMATION:
; APPLICANT: Flegel, Willy A.
; APPLICANT: DRK Blutspendedienst Baden-Wuerttemberg gGmbH
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES CORRELATED
; FILE REFERENCE: 12086-002001
; CURRENT APPLICATION NUMBER: US/09/600,714
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/EP98/08319
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Homosapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1251)
09-600714-2
MSKRYRSVRRCPLPMAILIEAAILLFFFTTHYDASLEDOKGLVASYOGDLYTMAAIGLGFITSSR
RHSWSSVAENLEMLALGVOMAILLDGFLSQFPGKAVITLFSIRLATMSALSVLISVDVGLGVNLAQLV
VMVLVEVTALGNLRMYISNIFNTDYMNMNMHIYFAAYFGLVAMCLPKPLPEGTEDKDOTATIPSLSAM
LGLFLFMMFPSPNSALLRSPIERKNNAVNTYAAVSVYTAISGSSLAHPQGIKSTYGHSAVLPGVA
VGTSCHLIPSPMLAMVIGIAGLISVGAAYLPGCCNRVYIGIPHSSIMGYNFSLGLLEIITVILVID
TVGAGNGMIGFOVLISIGELSLAIVIALTSGLLTGILLNKIKAPHEAKYFPDDQYFWKFPHLAVGFXI
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2001, 07:36:56 ; Search time 16.42 Seconds
(without alignments)
1939.157 Million cell updates/sec

Title: 09-600714-2

Perfect score: 2125
Sequence: 1 MSCKYPRSVGRCLPLCALTL.....AKYFDQYFWKPHLAVGF 418

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR-68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1981	93.2	417	2	152615 gene Rhd protein -
2	1976	93.0	417	2	A46368 erythrocyte membra
3	1863	87.7	417	2	I37005 Rhesus-like protei
4	1842	86.7	417	2	I37006 Rhesus-like protei
5	1840	86.6	417	2	I37004 Rhesus-like protei
6	1833	86.3	417	2	A30405 erythrocyte membra
7	1811	85.2	417	2	I37003 Rhesus-like protei
8	1803	84.8	417	2	I37075 Rhesus-like protei
9	1509	71.0	417	2	I84434 Rhesus-like protei
10	1411	66.4	354	2	S78480 Rhesus blood group
11	1294	60.9	354	2	I37053 Rhesus-like protei
12	1089.5	51.3	266	2	S78479 Rhesus blood group
13	1000.5	47.1	267	2	S78478 Rhesus blood group
14	547.5	25.8	409	2	S29124 membra glycoprot
15	412	19.4	463	2	T29442 hypothetical prote
16	361.5	17.0	457	2	T18673 RbPI-beta polypept
17	161	7.6	44	2	PC2033 RbPI-beta polypept
18	128	6.0	475	2	T01260 probable ammonium
19	127	6.0	458	2	C82232 proton/glutamate s
20	126	5.9	421	2	B84129 ammonium transport
21	125.5	5.9	468	2	A69468 ammonium transport
22	120	5.6	537	1	F70705 hypothetical prote
23	119.5	5.6	388	2	C64722 Na+/H+-exchanging
24	119.5	5.6	424	2	B69172 cationic amino aci
25	119.5	5.6	891	2	B82495 probable NADH dehy
26	119	5.6	881	2	S46633 probable membra
27	118	5.6	42	2	PC2032 RbPI-alpha polypep
28	117.5	5.5	388	2	A85482 Na+/H antiporter,
29	117.5	5.5	612	2	G64678 NADH dehydrogenase

30	113	5.3	653	2	D82352 Iron(III) ABC tran
31	113	5.3	782	2	T25925 hypothetical prote
32	112.5	5.3	419	2	D64142 hypothetical prote
33	112	5.3	385	2	G82789 permease Xf0589 (t
34	111.5	5.2	437	2	JC4988 high-affinity gluc
35	111.5	5.2	449	2	H69423 branched-chain am
36	111	5.2	660	2	A64739 ferichrome-iron t
37	110.5	5.2	612	2	E71839 NADH dehydrogenase
38	110	5.2	438	2	H86006 hypothetical prote
39	110	5.2	439	2	B75487 ammonium transport
40	110	5.2	507	2	T33024 hypothetical prote
41	109.5	5.2	429	2	F86021 arsenical pump mem
42	109.5	5.2	436	2	S47723 arsenite efflux pu
43	109.5	5.2	460	2	H83264 probable transport
44	109	5.1	363	2	A72771 hypothetical prote
45	109	5.1	461	2	D81418 probable transmem

ALIGNMENTS

RESULT 1
152615
gene Rhd protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 152615
R:Hang, C.H.; Reid, M.E.; Chen, Y.
Blood 86, 784-790, 1995
A:Title: Identification of a partial internal deletion in the RH locus causing the hu
A:Reference number: 152615; MUID:95329738
A:Accession: 152615
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:S78509; NID:g999309; PIDN:AAB34852.1; PID:g999310
A:Genetics:
A:Gene: Rhd

Query Match	Best Local Similarity	Score	Pred. No.	Length	DB 2;	Length	417;
Matches 396;	Conservative	4;	Mismatches 17;	Indels	0;	Gaps	0;
QY	1	MSCKYPRSVGRCLPLCALTLLEAALILFFFTTHDASLEDKGLVASYQVQDLTVMAAI	60				
DB	1	MSKYPKRSVRRCCLPCALTLLEAALILFFFTTHDASLEDKGLVASYQVQDLTVMAAI	60				
QY	61	GLGFTTSSFRHSSSVAFNFMALGVQWAILLDGFLSOPPSGKVITLFSIATWSA	120				
DB	61	GLGFTTSSFRHSSSVAFNFMALGVQWAILLDGFLSOPPSGKVITLFSIATWSA	120				
QY	121	LSVLISVDVAVGLVLAQLVVAVLEVVYDLGNLRVVISNIENPTYHNMHIIYFAAYFG	180				
DB	121	LSVLISVDVAVGLVLAQLVVAVLEVVYDLGNLRVVISNIENPTYHNMHIIYFAAYFG	180				
QY	181	LTVMACLPKPLPEGTEDDQATIPSLSAMGALFLMFTPSVNSALLRSTIERKNAVEN	240				
DB	181	LTVMACLPKPLPEGTEDDQATIPSLSAMGALFLMFTPSVNSALLRSTIERKNAVEN	240				
QY	241	TYVAVAASVTAIAIGSSLAHPQGRISKTYGSAVLPREGVAVATSCHLIPSPWMLAVIGLV	300				
DB	241	TYVAVAASVTAIAIGSSLAHPQGRISKTYGSAVLPREGVAVATSCHLIPSPWMLAVIGLV	300				
QY	301	AGLISVIGAKYLPGCCNVEIGIPHSSINGYNSLGLLEETITVLLVLDTVGANGMIG	360				
DB	301	AGLISVIGAKYLPGCCNVEIGIPHSSINGYNSLGLLEETITVLLVLDTVGANGMIG	360				
QY	361	FOVLLSTIGELSLAIVIALTSGLTALLINKITRAPHNAKFDQVFWKPHLAVGF	417				
DB	361	FOVLLSTIGELSLAIVIALTSGLTALLINKITRAPHNAKFDQVFWKPHLAVGF	417				

RESULT 2

A: Molecule type: protein Rhd - human
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Map position: 1p36.2-1p34
A: Cross-references: GDB:119551; OMIM:111680
C: Keywords: erythrocyte; surface antigen; transmembrane protein

Query Match 93.0%; Score 1976; DB 2; Length 417;
Best Local Similarity 95.0%; Pred. No. 1,7e-152;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60
DB 1 MSKPRSVGRCLPLMALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60

QY 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120
DB 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120

QY 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240
DB 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240

QY 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300
DB 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300

QY 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360
DB 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360

QY 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417
DB 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417

RESULT 3

A: Molecule type: protein - chimpanzee
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Map position: 1p36.2-1p34
A: Cross-references: GDB:119551; OMIM:111680
C: Keywords: erythrocyte; surface antigen; transmembrane protein

Query Match 93.0%; Score 1976; DB 2; Length 417;
Best Local Similarity 95.0%; Pred. No. 1,7e-152;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60
DB 1 MSKPRSVGRCLPLMALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60

QY 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120
DB 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120

QY 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240
DB 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240

QY 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300
DB 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300

QY 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360
DB 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360

QY 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417
DB 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417

RESULT 2

A: Molecule type: protein Rhd - human
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Map position: 1p36.2-1p34
A: Cross-references: GDB:119551; OMIM:111680
C: Keywords: erythrocyte; surface antigen; transmembrane protein

Query Match 93.0%; Score 1976; DB 2; Length 417;
Best Local Similarity 95.0%; Pred. No. 1,7e-152;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60
DB 1 MSKPRSVGRCLPLMALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60

QY 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120
DB 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120

QY 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240
DB 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240

QY 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300
DB 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300

QY 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360
DB 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360

QY 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417
DB 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417

RESULT 3

A: Molecule type: protein - chimpanzee
A: Residues: 401-407, 'DI' <SUY>
R: Westhoff, C.M.; Wylie, D.E.
Blood 83, 3098-3100, 1994
A: Title: Identification of a new Rhd-specific mRNA from K562 cells.
A: Reference number: 152589; PMID:94235883
A: Accession: 152589

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE2>
A: Cross-references: GB:S70174; NID:9546725; PIDN:AAB30756.1; PID:9546726
R: Suyama, K.; Lunn, R.; Haller, S.; Goldstein, J.
Blood 84, 1975-1981, 1994
A: Title: Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform in human
A: Reference number: 152596; PMID:94362249
A: Accession: 152596

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-217, 'M', 219-417 <RE3>
A: Cross-references: GB:S73913; NID:9693866; PIDN:AAB31911.1; PID:9693867
C: Genetics:
A: Gene: GDB:RHD; RH; RHD
A: Map position: 1p36.2-1p34
A: Cross-references: GDB:119551; OMIM:111680
C: Keywords: erythrocyte; surface antigen; transmembrane protein

Query Match 93.0%; Score 1976; DB 2; Length 417;
Best Local Similarity 95.0%; Pred. No. 1,7e-152;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKPRSVGRCLPLCALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60
DB 1 MSKPRSVGRCLPLMALTEALILLFYFTHYDASLEDDKGLVASYGODLTWMAI 60

QY 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120
DB 61 GLGFLTSSFRHSSSVAFNLMALGVQWAILLDGLSPFGKVVITLFTSLTMSA 120

QY 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 121 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180
DB 181 LSVLISVDALGVKNLAQVWVWLVETGLGNRMVSNFTNDYHNMHIVFAAYEG 180

QY 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240
DB 181 LVAWCLPKPLPGTEEDDQATPISLAMLGLFMTIFPSPNSALLRSPERKNAVEN 240

QY 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300
DB 241 TYAAVANSVYTAISGSSLAHPQKISTYGHSAVLPREGAVDTSCHLIPRPIVLGLV 300

QY 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360
DB 301 AGLSVYGAKYLPQCCNRVIGIPHSIMGYNFSILGLLEIYIVLLDVTGAGNGMIG 360

QY 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417
DB 361 FOYVLISIGELSLAIVALTSGLLTALLNLKIRKAPHEAKYFPDQYFWKPRHLAVGE 417

A:Residues: 1-417 <RES>
A:Cross-references: GB:I37050; NID:g606991; PIDN:AAA65624.1; PID:g606992

Query Match 87.7%; Score 1863; DB 2; Length 417;
Best Local Similarity 89.2%; Pred. No. 2,4e-143;
Matches 372; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

```

OY 1 MSCKYPSRSVRCPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSCKYPSRSVRCPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
OY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
OY 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
OY 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
OY 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
OY 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
OY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
OY 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
OY 361 FOVLISIGELSLAIVIALTSLGTLTGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 FOVLIRIGESLATTIATLSGLTLGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417

```

RESULT 4

Rhesus-like protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C/Accession: I37076
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colln, Y.; Ruffe,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085395
A:Accession: I37076
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37053; NID:g607011; PIDN:AAA65627.1; PID:g607012

Query Match 86.7%; Score 1842; DB 2; Length 417;
Best Local Similarity 89.0%; Pred. No. 1.2e-141;
Matches 371; Conservative 11; Mismatches 35; Indels 0; Gaps 0;

```

OY 1 MSCKYPSVGRCLPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSCKYPSVGRCLPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
OY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
OY 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
OY 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240

```

```

OY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
OY 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
OY 361 FOVLISIGELSLAIVIALTSLGTLTGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 FOVLIRIGESLATTIATLSGLTLGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417

```

RESULT 5

Rhesus-like protein - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C/Accession: I37004
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colln, Y.; Ruffe,
Biochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with t
A:Reference number: I37003; MUID:95085395
A:Accession: I37004
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:I37049; NID:g606989; PIDN:AAA65623.1; PID:g606990

Query Match 86.6%; Score 1840; DB 2; Length 417;
Best Local Similarity 88.7%; Pred. No. 1.8e-141;
Matches 370; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

```

OY 1 MSCKYPSVGRCLPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSCKYPSVGRCLPLCALTEALILLFYFPTHYDASLEDDKGLVASYOVGODLTWMAI 60
OY 61 GLGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GFGFLTSSFRHSSWSVAFLFMLALGVOMAILDGLFLOPPSGKVITTFISILATMSA 120
OY 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LSVLISDAVLGKYNLQOLVVMVLEVTDLGNLRMVISNIFNTDYHNMNMHIYFAAYFG 180
OY 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LTVAMCPLPKPECTEDNDQATIPSLSAMLGALFLMMPSPVNSALIRSPIERKNVFN 240
OY 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TYAVAVSVVTAISGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPWLPVLGLV 300
OY 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 AGLISVCGAKYLPCCNRRVLGIPHSSIMGYNFSLLGLEIITYVLVLDTVGANGMIG 360
OY 361 FOVLISIGELSLAIVIALTSLGTLTGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 FOVLIRIGESLATTIATLSGLTLGLLNKIKRAPHAKYFDQVFWKPEHLAVGF 417

```

RESULT 6

A30405
erythrocyte membrane protein Rh (unknown specificity) - human
N:Alternate names: blood group Rh antigen
C:Species: Homo sapiens (man)
C>Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 21-Jul-2000
C/Accession: A30405; S13060; I4273; S40515; S40516; S40517; S02085; S02087; A32509;
R:Cherif-Zahar, B.; Bloy, C.; Le Van Kim, C.; Blanchard, D.; Bailly, P.; Hermand, P.;
Proc. Natl. Acad. Sci. U.S.A. 87, 6243-6247, 1990

[illegible]

```

RESULT      8
I37075
Rhesus-like protein - gorilla
C:/Species: Gorilla gorilla (gorilla)
C:/Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C:/Accession: I37075
R:/Salvignol, I., Blancher, A., Calvas, P., Clayton, J., Socha, W.W., Colln, Y., Ruffie,
Biochem. Genet. 32, 201-221, 1994
A:/Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:/Reference number: I37003; MUID:95085595
A:/Accession: I37075
A:/Status: preliminary; translated from GB/EMBL/DBJ
A:/Molecule type: mRNA
A:/Residues: 1-417 <RES>
A:/Cross-references: GB:I37052; NID:9607009; PID:AAA56526.1; PID:9607010

```

Query Match	84.8%	Score 1803	DB 2	Length 417
Best Local Similarity	86.6%	Pred. No. 1.7e138		
Matches 361	Conservative 15	Mismatches 41	Indels 0	Gaps 0

Qy	1	MSCKPRSVGRCLEPCALTEEAALILFEYFPTHYDASLEOQKGLVASYQGDILTMAAI	60
Db	1	MSCKPRSVRCCELEPCALTEEAALILFEYFPTHYDASLEOQKGLVASYQGDILTMAAI	60
Qy	61	GIAGLTSSFRRHSMSSVAFLNFMALGVOMAILLDGELSOPGSGKVITTEFSIWLATMSA	120
Db	61	GFGLTSSFRRHSMSSVAFLNFMALGVOMAILLDGELSDOPPEKKVITTEFSIRLATMSA	120
Qy	121	LSVLISVDALVAKYNLAQLVVVLVEVTDGCLNRMVTSINFTNDYHNMNMHIYEAAYFG	180
Db	121	LSVLISAGAVLGVNVLVQLVVVLVEVYALGTMRVTSINFTNDYHNMNMHIYEAAYFG	180
Qy	181	LTVMACRPLPDIKEDDDQATIPSLSAMIGALFLMMFPRSVSALLRSPRIEKNVFN	240
Db	181	LTVMACRPLPDIKEDDDQATIPSLSAMIGLFLMMFMPSFSSALLRSPRIEKNVFN	240
Qy	241	TYVAANVSVAIISGSSLIAPHOGKISKRTYGHSAVLPEGVANDTSCHLITSPMLDpLVGLV	300
Db	241	TYVALANSVVAIISVSSLIAPHOGKINMTYHMNAVLAGVAVGTSCHLITSPMLAMVGLV	300
Qy	301	AGLISVIGAKYLPGCCNKNKVLGIPHSSIMGVNFSLLGLLEIYIVLVLDTLVGAGNGMIG	360
Db	301	AGLISIGAKCLPGCCNKNKVLGIDHSSVMHYNFSLLGLLEIITYIVLAVLHVGAGNGMG	360
Qy	361	FOVLLSTBELSLAIVIALTSLTALLNLKITAPHEAKYFDDQVYWKPRPHLAVGF	417
Db	361	FOVAVSTBELSLAIVAVTSLLGLLNLKITAPHAAKYFDDQVYWKPRPHLAVGF	417

RESULT 9
184434
Rhesus-like protein - crab-eating macaque
C:Species: Macaca fascicularis (Crab-eating macaque)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C:Accession: 184434
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffie, Biochem. Genet. 32, 201-221, 1994
A:File: Molecular genetics of chimpanzee Rh-related genes: their relationship with the A:Reference number:137003; MUID:95085595

A:Accession: I84434
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:IJ37054; NID:9609515; PIDDN:AA65628.1; PID:g609516

Query Match	71.0%	Score 1509;	DB 2;	Length 417;
Best Local Similarity	74.3%	Pred. No. 1.1e-114;		
Matches 310;	Conservative 31;	Mismatches 76;	Indels 0;	Gaps 0;

```

0Y      1 MSCKPRSVGVCPLCALTELEAALILFEYFETHDASLEPOKGVASVOYQDUTMAAI 60
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      1 MSSKIPMSVRCDEPMALITELALILFEFFTYTIDASLEQKGVASVOYQDUTMAAV 60
0Y      61 GIGELTSSFFRRSHSSVAFLFMALGVOMAILDGLFSOPSPGKVITLFTSLWTMSA 120
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      61 GIGFTTSLRNRNMSSSVAFNLFLLATGVOMAILDGLFSOPGKVYIKLFTSLRTRST 120
0Y      121 LSVLISDVAVLGKYNLAOLVYVYVEVYTDIGNLRMVTSINFNDYHNNMMHLYFAATG 180
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      121 TSMLSMNAVLGKYNLAOLVYMELEVETLVGTMYRIVYINFEKIDYGNMMHIFVFAATG 180
0Y      181 LTVAMCDEKPLPECTEGEDDOATIPESLAMIAGLFEMFEPYNSALLRSPERKNAVEN 240
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      181 LTVAMCDEKPLPKGTEDDYQTTTSPSLFAMIGTFLMFMWPTNSALLRPERKNAVES 240
0Y      241 TYAAVAVSVAIVAISSSLAHPOGKISKTYGHSALVPEGVAVDTSCHLITSPMIPVILGV 300
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      241 TYTAAVSAVTAIVSYSSLAHPOKRIKMTYMPRNAGLAGVAVASCHYIHSPTIAMVLGV 300
0Y      301 AGLISVIGAKYLLPGCCNFRVLGIPHSSIMGVNSLLGLLEETIYVLLVDTVGAGNGMTG 360
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      301 AGLISFGAKCLPCVCFNRVLGIEHSHSMHTYEGPLALGETIYVILMALRVWASSNMIG 360
0Y      361 FOVLSTIGELSIAIVIALTSLSGLTALLNLKLTAAPEAKYFPDDQVPMKPRHLAVGF 417
      1 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      361 FOVLSTIGELSIAAMSTISOLLGLINKIMKGPVAVFPDDQAFWEPKHLAVGF 417

```

RESULT 10
S78480
Rhesus blood group antigen-like protein isoform 4 - human
C:Species: Homo sapiens (man)
C:Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
R:Accession: S78480; S26560
R:Colln, Y.
submitted to the EMBL Data Library, November 1991
A:Reference number: S78478
A:Accession: S78480
A:Molecule type: mRNA
A:Residues: 1-354 <COL>
A:Cross-references: EMBL:X63098; NID:936019; PIDN:CAAA4812.1; PID:936020
R:Le Van Kim, C.; Cherif-Zahar, B.; Raynal, V.; Mourou, I.; Lopez, M.; Cartron, J.P.;
Blood 80, 1074-1078, 1992
A>Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A:Reference number: S26560; MUID:92360855
A:Accession: S26560
A:Molecule type: mRNA
A:Residues: 160-354 <LEV>
A:Cross-references: EMBL:X63098
A:Experimental source: tissue-type bone marrow
C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

Query Match	66.4%	Score 1411	DB 2	Length 354
Best Local Similarity	90.1%	Pred No. 8e-107		
Matches	282	Conservative	10	Mismatches 21; Indels 0; Gaps 0

QY	1	MSCKYPRSVRQCLPCALTLLEAALLILEFFPHYDASLEDDQGLVASVYGVGGDLYMAAI	60
Db	1	MSCKYPRSVRQCLPCALTLLEAALLILEFFPHYDASLEDDQGLVASVYGVGGDLYMAAL	60

QY 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
|||||
Db 61 CLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
QY 121 LSVLISDAVLGKYNALQVVMVLEVTGLGNLRMVISNIFNTDYHNMHHIYFAAYFG 180
|||||
Db 121 MSVLISGAVLGKYNALQVVMVLEVTALGTLRMVISNIFNTDYHNMHHIYFAAYFG 180
QY 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMLGALFLMWFPSVNSALLRSPIERKNAVEN 240
|||||
Db 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMLGALFLMWFPSVNSALLRSPIERKNAVEN 240
QY 241 TYVAVASVVTALISGSSLAHPQKISKTGYHSAVLPQGVAVDTSCHLIPSPWLIVGLV 300
|||||
Db 241 TYVAVASVVTALISGSSLAHPQKISMTYHSAVLAAGVAVGTSCHLIPSPWLIVGLV 300
QY 301 AGLISVFGAKYLP 313
|||||
Db 301 AGLISVFGAKYLP 313

RESULT 11

I37053
Rhesus-like protein - pileated gibbon
C:Species: Hylobates pileatus (pileated gibbon)
C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 05-Nov-1999
C:Accession: I37053
R:Salvagno, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Collin, Y.; Ruffie, Blochem. Genet. 32, 201-221, 1994
A:Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the A:Reference number: I37003; MUID:95085595
A:Accession: I37053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: GB:I37051; NID:9607007; PIDN:AAA65625.1; PID:9607008

Query Match 60.9%; Score 1294; DB 2; Length 354;
Best Local Similarity 82.7%; Pred. No. 2.4e-97;
Matches 259; Conservative 20; Mismatches 34; Indels 0; Gaps 0;
QY 1 MSCKYPRSVGRCPLCALTEALILLFYFTHYDASLEDDKGLVASVOYGOGLTVMAAI 60
|||
Db 1 MSCKYPRSVRCPLCALTEALILLFYFTHYDASLEDDKGLVATYVOGDLTVMAAL 60
QY 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
|||||
Db 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
QY 121 LSVLISDAVLGKYNALQVVMVLEVTGLGNLRMVISNIFNTDYHNMHHIYFAAYFG 180
|||||
Db 121 MSVLISGAVLGKYNALQVVMVLEVTALGTLRMVISNIFNTDYHNMHHIYFAAYFG 180
QY 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMLGALFLMWFPSVNSALLRSPIERKNAVEN 240
|||||
Db 181 LTVAMCLPKPEGTEDNEQGLANSPSLSAMLGALFLMWFPSVNSALLRSPIERKNAVEN 240
QY 241 TYVAVASVVTALISGSSLAHPQKISKTGYHSAVLPQGVAVDTSCHLIPSPWLIVGLV 300
|||||
Db 241 TYVAVASVVTALISVSSLAHPQKINMTYHSAVLAAGVAVGTSCHLIPSPWLIVGLV 300
QY 301 AGLISVFGAKYLP 313
|||||
Db 301 AGLISVFGAKYLP 313

RESULT 12

S78479
Rhesus blood group antigen-like protein isoform VIII - human
C:Species: Homo sapiens (man)
C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
C:Accession: S78479; S26563

R:Collin, Y.
submitted to the EMBL Data Library, November 1991
A:Reference number: S78478
A:Accession: S78479
A:Molecule type: mRNA
A:Residues: 1-266 <COL>
A:Cross-references: EMBL:X63096; NID:936044; PIDN:CAA44810.1; PID:936045
R:Le Van Kim, C.; Cherif-Zahar, B.; Raynal, V.; Mouro, I.; Lopez, M.; Carttron, J.P.; Blood 80, 1074-1078, 1992
A:Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A:Reference number: S26560; MUID:92360853
A:Accession: S26563
A:Molecule type: mRNA
A:Residues: 160-266 <LEV>
A:Cross-references: EMBL:X63096
A:Experimental source: tissue-type bone marrow
C:Keywords: alternative splicing; erythrocyte; surface antigen; transmembrane protein

Query Match 51.3%; Score 1089.5; DB 2; Length 266;
Best Local Similarity 58.0%; Pred. No. 6.2e-81;
Matches 242; Conservative 3; Mismatches 21; Indels 151; Gaps 1;

QY 1 MSCKYPRSVGRCPLCALTEALILLFYFTHYDASLEDDKGLVASVOYGOGLTVMAAI 60
|||
Db 1 MSCKYPRSVRCPLCALTEALILLFYFTHYDASLEDDKGLVASVOYGOGLTVMAAI 60
QY 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
|||||
Db 61 GLGFLTSSFRHSSVAFNLFMLALGVOMAILLDGFLSQFPSPKVVITLFSILATMSA 120
QY 121 LSVLISDAVLGKYNALQVVMVLEVTGLGNLRMVISNIFNTDYHNMHHIYFAAYFG 180
|||||
Db 121 MSVLISGAVLGKYNALQVVMVLEVTALGTLRMVISNIFNV----- 163
QY 181 LTVAMCLPKPLPEGTEDNDQATIPSLSAMLGALFLMWFPSVNSALLRSPIERKNAVEN 240
|||||
Db 164 ----- 163
QY 241 TYVAVASVVTALISGSSLAHPQKISKTGYHSAVLPQGVAVDTSCHLIPSPWLIVGLV 300
|||||
Db 164 ----- 163
QY 301 AGLISVFGAKYLPCCNRVIGIPHSIMGYNFSLLGLEIYVLLVDVTCAGNGMIG 360
|||||
Db 164 -----CCNRVIGIHHSVMHSFSLGLGELTYVLLVLTVMNGNMIG 209
QY 361 FOYLLSIGELSLAIVIALTSGLLTALLNLKIRKAPHEAKYFPDQVFKRPHLAVGF 417
|||||
Db 210 FOYLLSIGELSLAIVIALTSGLLTALLNLKIRKAPHAKYFPDQVFKRPHLAVGF 266

RESULT 13

S78478
Rhesus blood group antigen-like protein isoform VI - human
C:Species: Homo sapiens (man)
C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 05-Nov-1999
C:Accession: S78478; S26562
R:Collin, Y.
submitted to the EMBL Data Library, November 1991
A:Reference number: S78478
A:Accession: S78478
A:Molecule type: mRNA
A:Residues: 1-267 <COL>
A:Cross-references: EMBL:X63095; NID:936042; PIDN:CAA44809.1; PID:936043
R:Le Van Kim, C.; Cherif-Zahar, B.; Raynal, V.; Mouro, I.; Lopez, M.; Carttron, J.P.; Blood 80, 1074-1078, 1992
A:Title: Multiple Rh messenger RNA isoforms are produced by alternative splicing.
A:Reference number: S26560; MUID:92360855
A:Accession: S26562
A:Molecule type: mRNA
A:Residues: 160-267 <LEV>
A:Cross-references: EMBL:X63095

Db 362 QYDKIYPGMARGEDRTRMFDEKTOALNQLMAIGLVFLASTV---SGYLTGILLKTIWDQ 418
QY 396 PHEAKYFDDQYFWKFP 411
: :1: | : : : |
Db 419 VRDDEYYADGSDYFETP 434

Search completed: September 12, 2001, 07:39:35
Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2001, 07:39:11 ; Search time 9.37 Seconds

(without alignments)
1528.153 Million cell updates/sec

Title: 09-600714-2

Perfect score: 2125

Sequence: 1 MSCKYPSRVGRCLPLCALTL.....AKYFDQYFWKPEHLAVGEF 418

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1975	92.9	416	1	RHD_HUMAN
2	1858	87.4	416	1	RHLR_PANTR
3	1837	86.4	416	1	RHLD_GORGO
4	1835	86.4	416	1	RHLA_PANTR
5	1828	85.0	416	1	RHCE_HUMAN
6	1806	85.0	416	1	RHFE_PANTR
7	1798	84.6	416	1	RHLC_GORGO
8	1519	71.5	416	1	RHL_MACMU
9	1504	70.8	416	1	RHL_MACFA
10	1289	60.7	409	1	RHAG_HUMAN
11	547.5	25.8	409	1	RHAG_HUMAN
12	119.5	5.6	388	1	NHAA_ECOLI
13	112.5	5.6	881	1	YJTB_YEAST
14	111.5	5.3	419	1	Y092_HAEIN
15	111.5	5.2	437	1	GNTB_ECOLI
16	111	5.2	660	1	FHUB_ECOLI
17	109.5	5.2	429	1	ARSB_ECOLI
18	109.5	5.2	429	1	ARSB_ECOLI
19	109	5.1	613	1	NUSM_POLOR
20	107.5	5.1	403	1	YDHC_ECOLI
21	107.5	5.1	446	1	GNTU_ECOLI
22	107	5.0	643	1	NUSM_ASTPE
23	106.5	5.0	429	1	ARSB_YEREN
24	106.5	5.0	507	1	Y108_SYNY3
25	106	5.0	401	1	CHRA_AICEU
26	104	4.9	603	1	NUSM_PONPA
27	103.5	4.9	451	1	GNTP_ZYMO
28	103.5	4.9	547	1	NUSM_ASCSU
29	103.5	4.9	788	1	COXM_SULAC
30	102.5	4.8	1232	1	B3A3_HUMAN
31	102	4.8	449	1	YJCD_ECOLI
32	100.5	4.7	491	1	NUOM_RICPR
33	100	4.7	383	1	NAPA_ENTHR

34	100	4.7	416	1	YBDA_ECOLI	P24077	escherichia
35	100	4.7	452	1	NU4M_BRAFL	074423	branchiosto
36	100	4.7	452	1	NU4M_BRAFL	079421	branchiosto
37	100	4.7	887	1	Y277_MYCPN	P75387	mycoplasma
38	100	4.7	982	1	YS96_CABEL	009965	caenorhabdi
39	99.5	4.7	594	1	COXI_PHYPO	007434	physarum po
40	99	4.7	603	1	NUSM_PANPA	P03916	pan paniscu
41	98.5	4.6	448	1	GNTP_BACSU	P12012	bacillus su
42	98.5	4.6	509	1	NUOM_ECOLI	P31978	escherichia
43	98	4.6	338	1	YFJY_ECOLI	P52137	escherichia
44	98	4.6	347	1	NU2M_DIDMA	P41305	didelphis m
45	98	4.6	442	1	YCDG_ECOLI	P75892	escherichia

ALIGNMENTS

RESULT 1

RHD_HUMAN STANDARD; PRT; 416 AA.

AC 002161; 002162; 007618; 016355; 016147; 016235;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BLOOD GROUP RH(D) POLYPEPTIDE (RHESUS D ANTIGEN) (RHXIII) (RH

DE POLYPEPTIDE 2) (RHPHII).

GN RHD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP TISSUE=Bone marrow;

RC MEDLINE=93066356; PubMed=1438298;

RA le van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,

RA Carton J.-P., Collin Y.;

RT "Molecular cloning and primary structure of the human blood group Rhd

RT polypeptide.";

RT Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).

RN [2]

RP TISSUE=FROM N.A.

RC MEDLINE=93320449; PubMed=8329718;

RA Arce M.A., Thompson E.S., Wagner S., Coyne K.E., Ferdman B.A.,

RA Lublin D.M.;

RT "Molecular cloning of Rhd cDNA derived from a gene present in Rhd-

RT positive, but not Rhd-negative individuals.";

RT Blood 82:651-655(1993).

RN [4]

RP TISSUE=FROM N.A.

RC MEDLINE=93216282; PubMed=7916743;

RA Kojil E., Umenishi F., Iwamoto S., Ikemoto S.;

RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated

RT with the Rh blood group system.";

RT Hum. Genet. 91:157-162(1993).

RN [5]

RP TISSUE=FROM N.A.

RC MEDLINE=95329738; PubMed=7606008;

RA Huang C.H., Reid M.E., Chen Y.;

RT "Identification of a partial internal deletion in the RH locus

RT causing the human erythrocyte D-phenotype.";

RT Blood 86:784-790(1995).

RN [6]

RP TISSUE=FROM N.A. (SHORT FORM 1).

```

RX MEDLINE=94235883; PubMed=8180407;
RA Westhoff C.M., Wylie D.E.;
RT "Identification of a new RhD-specific mRNA from K562 cells.";
RL Blood 83:3098-3100(1994).
RP
RX SEQUENCE FROM N.A. (SHORT FORM 2).
RX MEDLINE=94362249; PubMed=8080999;
RA Suyama K., Lunn R., Haller S., Goldstein J.;
RT "Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform
in human erythroleukemic K562 cells.";
RL Blood 84:1975-1981(1994).
RN
RP VARIANT BLOOD GROUP TAR.
RX MEDLINE=95259709; PubMed=7741145;
RA Rouillac C., le van Kim C., Beolet M., Cartion J.-P., Colin Y.;
RT "Leu10Pro substitution in the RhD polypeptide is responsible for the
DVI1 category blood group phenotype.";
RL Am. J. Hematol. 49:87-88(1995).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING
CC ERYTHROID CHARACTERS.
CC -1- POLYMORPHISM: RHD AND RHCE ARE RESPONSIBLE FOR THE RH BLOOD GROUP
CC SYSTEM. THE MOLECULAR BASIS OF THE TAR-RH40 BLOOD GROUP ANTIGEN
CC IS A POLYMORPHISM IN POSITION 109.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHCE.
CC
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X63097; CAA44811.1; -
DR EMBL: X63094; CAA44808.1; -
DR EMBL: L08429; AAA02679.1; -
DR EMBL: S57971; AAB36081.1; -
DR EMBL: S78509; AAB34852.1; -
DR EMBL: S70174; AAB30756.1; -
DR EMBL: S73913; AAB31911.1; -
DR PIR: S26564; S26564.
DR MIM: 111680; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
KM Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing;
KM Polymorphism.
FT INIT_MET 0
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT VARSPLIC 313 408 MISSING (IN SHORT ISOFORM 1).
FT VARSPLIC 315 416 C -> S (IN SHORT ISOFORM 2).
FT VARSPLIC 316 416 MISSING (IN SHORT ISOFORM 2).
FT VARIANT 109 109 L -> P (IN TAR ANTIGEN).
FT VARIANT 217 217 M -> I.
FT VARIANT 217 217 /FTID=VAR_006920.
FT CONFLICT 15 15 W -> C (IN REF. 5).

```

```

FT CONFLICT 38 38 E -> G (IN REF. 4).
FT CONFLICT 102 102 S -> P (IN REF. 4).
FT CONFLICT 126 126 V -> A (IN REF. 4).
FT CONFLICT 173 173 V -> M (IN REF. 5).
FT CONFLICT 181 181 S -> T (IN REF. 4).
FT CONFLICT 313 313 G -> V (IN REF. 4 AND 7).
FT CONFLICT 322 322 P -> H (IN REF. 4).
FT CONFLICT 397 397 E -> V (IN REF. 5).
SQ SEQUENCE 416 AA; 45049 MM; 9136DPLA37D76B1B CRC64;

Query Match          92.9%; Score 1975; DB 1; Length 416;
Best Local Similarity 95.2%; Pred. No. 1.2e-136;
Matches 396; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 2 SKYPRSVGCLPLCALTEALALLLFYFTTHDASLEDDKGLVASYQVQODLTWMAIG 61
   1 SSKYPRSVRCCLPLMALTEALALLLFYFTTHDASLEDDKGLVASYQVQODLTWMAIG 60
DB
QY 62 LGFTSSFRHSSSVAFNLFMLALGVQWAILLDGFLSOPPSGVVTLTSLATWSAL 121
   61 LGFTSSFRHSSSVAFNLFMLALGVQWAILLDGFLSOPPSGVVTLTSLATWSAL 120
DB
QY 122 SVLISDAVLGKYNLAQLVYVAVLVETDLGNLRVYVSNINFTDYHNMNHIYFAAYFGL 181
   121 SVLISDAVLGKYNLAQLVYVAVLVETDLGNLRVYVSNINFTDYHNMNHIYFAAYFGL 180
DB
QY 182 TVAMCLPKPLPEGETEDDQATATPSLSAMGALFLMFFSVSALLRSIERKNVNT 241
   181 TVAMCLPKPLPEGETEDDQATATPSLSAMGALFLMFFSVSALLRSIERKNVNT 240
DB
QY 242 YVAVAVSVVAISGSSSLAHPOGKISKTYGHSAYLPGCAVATSCHLIPSPWLPVLGLVA 301
   241 YVAVAVSVVAISGSSSLAHPOGKISKTYGHSAYLPGCAVATSCHLIPSPWLPVLGLVA 300
DB
QY 302 GLISVTCATKLPGCCNRYLGIPIHSSITGVNFSLGLGLEIITYLVLLDTVAGNGMIGF 361
   301 GLISVTCATKLPGCCNRYLGIPIHSSITGVNFSLGLGLEIITYLVLLDTVAGNGMIGF 360
DB
QY 362 QVLLSTIGELSLAIVALTSGLLTALNLKIRKAPHEAKYEDDOVFKPFLHVG 417
   361 QVLLSTIGELSLAIVALTSGLLTALNLKIRKAPHEAKYEDDOVFKPFLHVG 416
DB

RESULT 2
RHLR_PANTR STANDARD; PRT; 416 AA.
AC Q28814;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN IIR (RHESUS-LIKE PROTEIN IIR).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OC NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human Rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: L37050; AAA65624.1; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium transp. 1.
DR PRINTS: PR00342; RHESUSRD.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0
FT TRANSMEM 11 31
FT TRANSMEM 43 63
FT TRANSMEM 76 96
FT TRANSMEM 124 144
FT TRANSMEM 171 191
FT TRANSMEM 202 222
FT TRANSMEM 237 257
FT TRANSMEM 264 284
FT TRANSMEM 286 306
FT TRANSMEM 330 350
FT TRANSMEM 357 377
SQ SEQUENCE 416 AA; 45299 MM; C055D7CC8B4A0420 CRC64;
BY SIMILARITY.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.

Query Match 87.4%; Score 1858; DB 1; Length 416;
Best Local Similarity 89.2%; Pred. No. 3.8e-128;
Matches 371; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

QY 2 SCKYPRSVGRCLPLCALTLLEAAILLFFFTTHYDASLEDDKGLVASYOVGODLVMAIG 61
DB 1 SSKYPRSVGRCLPLCALTLLEAAILLFFFTTHYDASLEDDKGLVASYOVGODLVMAIG 60
QY 62 LGFTSSFRHRHSSVAFNFMALGVOMAILLDGFLSOPPGKVVITLTSIRLATMSAL 121
DB 61 FGFLTSSFRHRHSSVAFNFMALGVOMAILLDGFLSOPPGKVVITLTSIRLATMSAL 120
QY 122 SVLISVDVAVLGKVNLAOLVYVWVLEVTGLNLRVVISINFTDQVHMMMHIIYFAAYEGL 181
DB 121 SVLISVDVAVLGKVNLAOLVYVWVLEVTGLNLRVVISINFTDQVHMMMHIIYFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDDQATIPSLSAMIGALFLMFPVSNSALLRSPERKNAVFNT 241
DB 181 SVAMCLPKPLPEGETEDDQATIPSLSAMIGALFLMFPVSNSALLRSPERKNAVFNT 240
QY 242 YVAAVSVVTAISGSSLAHPQGIKSTYGHSAVLPBGVAVGTSCHLIPSPWLAVGLVA 301
DB 241 YVAAVSVVTAISGSSLAHPQGIKSTYGHSAVLPBGVAVGTSCHLIPSPWLAVGLVA 300
QY 302 GLISVCAKYLPGCCNNVLGIPHSIMGYNPSILGLLEIITLYLVLDVYGAAGNGMIGF 361
DB 301 GLISVCAKYLPGCCNNVLGIPHSIMGYNPSILGLLEIITLYLVLDVYGAAGNGMIGF 360
QY 362 QVLSIGELSLAIVIALTSGLLTALLNLKIRKAPHEAKFYDDQVFKFPHLAVGF 417
DB 361 QVLSIGELSLAIVIALTSGLLTALLNLKIRKAPHEAKFYDDQVFKFPHLAVGF 416

RESULT 3
RHLD_GORGO STANDARD; PRT; 416 AA.
AC 028427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE RHESUS-LIKE PROTEIN.
OS Gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE RH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: L37053; AAA65627.1; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium transp. 1.
DR PRINTS: PR00342; RHESUSRD.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0
FT TRANSMEM 11 31
FT TRANSMEM 43 63
FT TRANSMEM 76 96
FT TRANSMEM 124 144
FT TRANSMEM 171 191
FT TRANSMEM 202 222
FT TRANSMEM 237 257
FT TRANSMEM 264 284
FT TRANSMEM 286 306
FT TRANSMEM 330 350
FT TRANSMEM 357 377
SQ SEQUENCE 416 AA; 44838 MM; 398877FB7BEB802 CRC64;
BY SIMILARITY.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.

Query Match 86.4%; Score 1837; DB 1; Length 416;
Best Local Similarity 88.9%; Pred. No. 1.3e-126;
Matches 370; Conservative 11; Mismatches 35; Indels 0; Gaps 0;

QY 2 SCKYPRSVGRCLPLCALTLLEAAILLFFFTTHYDASLEDDKGLVASYOVGODLVMAIG 61
DB 1 SSKYPRSVGRCLPLCALTLLEAAILLFFFTTHYDASLEDDKGLVASYOVGODLVMAIG 60
QY 62 LGFTSSFRHRHSSVAFNFMALGVOMAILLDGFLSOPPGKVVITLTSIRLATMSAL 121
DB 61 FGFLTSSFRHRHSSVAFNFMALGVOMAILLDGFLSOPPGKVVITLTSIRLATMSAL 120
QY 122 SVLISVDVAVLGKVNLAOLVYVWVLEVTGLNLRVVISINFTDQVHMMMHIIYFAAYEGL 181
DB 121 SVLISVDVAVLGKVNLAOLVYVWVLEVTGLNLRVVISINFTDQVHMMMHIIYFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDDQATIPSLSAMIGALFLMFPVSNSALLRSPERKNAVFNT 241
DB 181 SVAMCLPKPLPEGETEDDQATIPSLSAMIGALFLMFPVSNSALLRSPERKNAVFNT 240
QY 242 YVAAVSVVTAISGSSLAHPQGIKSTYGHSAVLPBGVAVGTSCHLIPSPWLAVGLVA 301
DB 241 YVAAVSVVTAISGSSLAHPQGIKSTYGHSAVLPBGVAVGTSCHLIPSPWLAVGLVA 300
QY 302 GLISVCAKYLPGCCNNVLGIPHSIMGYNPSILGLLEIITLYLVLDVYGAAGNGMIGF 361
DB 301 GLISVCAKYLPGCCNNVLGIPHSIMGYNPSILGLLEIITLYLVLDVYGAAGNGMIGF 360
QY 362 QVLSIGELSLAIVIALTSGLLTALLNLKIRKAPHEAKFYDDQVFKFPHLAVGF 417
DB 361 QVLSIGELSLAIVIALTSGLLTALLNLKIRKAPHEAKFYDDQVFKFPHLAVGF 416

Db 361 QVLISGELSLAIVALTSLGLTGLLNLKIKWAPHAAYFDQVWKPFLAVGF 416

RESULT 4

RHIA_PANTR STANDARD: PRT: 416 AA.

ID RHIA_PANTR

AC Q28813;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE RH-LIKE PROTEIN IA (RHESUS-LIKE PROTEIN IA).

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

OC NCBI_TaxID=9598;

OX

RN

RP SEQUENCE FROM N.A.

RC TISSUE-BONE marrow;

RX MEDLINE=95085595; PubMed-7993375;

RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W., Collin Y., Ruffie J.;

RT "Molecular genetics of chimpanzee Rh-related genes: their relationship with the R-C-E-F blood group system, the chimpanzee counterpart of the human Rhesus system.";

RL Biochem. Genet. 32:201-221(1994).

CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L37049; AAA5623.1; -

DR InterPro: IPR001905; -

DR InterPro: IPR002229; -

DR Pfam: PF00909; Ammonium_transp; 1.

DR PRINTS: PR00342; RHESUSRD.

DR Erythrocyte; Transmembrane.

KW

FT INIT MET 0 0 BY SIMILARITY.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 43 63 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 124 144 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 237 257 POTENTIAL.

FT TRANSMEM 264 284 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 330 350 POTENTIAL.

FT TRANSMEM 357 377 POTENTIAL.

SO SEQUENCE 416 AA; 45461 MW; B0B566734DB5E14D CRC64;

Query Match 86.4%; Score 1835; DB 1; Length 416;

Best Local Similarity 88.7%; Pred. No. 1,8e-126;

Matches 369; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 SCYPRSVGCLPLCALTEALILFFFTTHDASLEDDKGLVASVQVQODLTVAATG 61

DB 1 SSRYPRSVRCLEPLCALTEALILFFFTQYDASLEDDKGLVASVQVQODLTVAATG 60

QY 62 LGFLTSFRSHSSVAFNFLMALGYOMAILLDGFLSOPPSGVVTLTFSIWLATMSAL 121

DB 61 FGLTSSFRSHSSVAFLMALGYOMAILLDGFLSOPPSGVVTLTFSIWLATMSAL 120

QY 122 SVLISDAVLGVNLAQVVMVLEVTDLGNLRMVISNIFNTDYNHNMHIYFAAYFGL 181

DB 121 SVLISDAVLGVNLAQVVMVLEVTALGTVMVISNIFNTDYNHNMHIYFAAYFGL 180

QY 182 TVAMCLPKPLPECTEDNDQATIPSLAMGALFLMMFPPSVNSALLRSPTEKKNVFNPT 241

DB 181 SVAMCLPKPLPECTEDNDQATIPSLAMGALFLMMFPPSVNSALLRSPTEKKNVFNPT 240

QY 242 YVAAVSVYTAISGSSLAHPDCKISKTYGSHAVLPQGVAVDTSCHLIPSPMLPIVGLVA 301

DB 241 YVAAVSVYTAISGSSLAHPDCKISKTYGSHAVLPQGVAVDTSCHLIPSPMLPIVGLVA 300

QY 302 GLISVYGAKYLPCCCNRVGLGIPHSSTMGVNFSLGLLEIYIVLVLDTVGAGNMIGF 361

DB 301 GLISVYGAKYLPCCCNRVGLGIPHSSTMGVNFSLGLLEIYIVLVLDTVGAGNMIGF 360

QY 362 QVLISGELSLAIVALTSLGLTGLLNLKIKWAPHAAYFDQVWKPFLAVGF 417

DB 361 QVLIRIGERSLATITALTSLGLTGLLNLKIKWAPHAAYFDQVWKPFLAVGF 416

RESULT 5

RHCE_HUMAN STANDARD: PRT: 416 AA.

ID RHCE_HUMAN

AC P18577; Q02163; Q02164; Q02165; Q16160;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BLOOD GROUP RH(C) POLYPEPTIDE (RHESUS C/E ANTIGENS) (RH30A) (RH1XB) (RH POLYPEPTIDE 1) (RHPI).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OC NCBI_TaxID=9606;

RN

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE-BONE marrow;

RX MEDLINE=90349591; PubMed-1696722;

RA Cherif-Zahar B., Bloy C., Le van Kim C., Blanchard D., Bailly P., Hernand P., Salmon C., Carton J.-P., Collin Y.;

RT "Molecular cloning and protein structure of a human blood group Rh polypeptide.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247(1990).

RN

RP SEQUENCE FROM N.A.

RX MEDLINE=91058522; PubMed-2123099;

RA Arent N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;

RT "cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rhesus)-blood-group-antigen expression.";

RL Biochem. J. 271:821-825(1990).

RN

RP SEQUENCE FROM N.A.

RX MEDLINE=93216282; PubMed-7916743;

RA Kajli E., Umenishi F., Iwamoto S., Ikemoto S.;

RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated with the Rh blood group system.";

RL Hum. Genet. 91:157-162(1993).

RN

RP SEQUENCE FROM N.A. (FORMS RHIV AND RHVIII).

RC TISSUE-BONE marrow;

RX MEDLINE=92360855; PubMed-1379850;

RA le van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M., Carton J.-P., Collin Y.;

RT "Multiple Rh messenger RNA isoforms are produced by alternative splicing.";

RL Blood 80:1074-1078(1992).

RN

RP SEQUENCE OF 1-32.

RX MEDLINE=89134163; PubMed-3146980;

RA Arent N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J., Kumpel B.;

RT "Protein-sequence studies on Rh-related polypeptides suggest the presence of at least two groups of proteins which associate in the human red-cell membrane.";

RL Biochem. J. 256:1043-1046(1988).
 RN [6]
 RN SEQUENCE OF 1-16.
 RX MEDLINE-88294325; PubMed-3135863;
 RA Bloy C., Blanchard D., Dahr W., Beyreuther K., Salmon C.,
 RA Cartton J.-P.;
 RT "Determination of the N-terminal sequence of human red cell Rh(D)
 RT polypeptide and demonstration that the Rh(D), (c), and (E) antigens
 RT are carried by distinct polypeptide chains.";
 RL Blood 72:661-666(1988).
 RN [7]
 RN SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE-94245182; PubMed-8188244;
 RA Cherif-Zahar B., Je van Kim C., Rouillac C., Raynal V., Cartton J.-P.,
 RA Colin Y.;
 RT "Organisation of the gene (RHCE) encoding the human blood group
 RT RHCE antigen and characterization of the promoter region.";
 RL Genomics 19:68-74(1994).
 RN [8]
 RN VARIANTS BLOOD GROUP C AND E.
 RX MEDLINE-94035121; PubMed-8220426;
 RA Mouro I., Colin Y., Cherif-Zahar B., Cartton J.-P., Je van Kim C.;
 RT "Molecular genetic basis of the human Rhesus blood group system.";
 RL Nat. Genet. 5:62-65(1993).
 CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
 CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING
 CC ERYTHROID CHARACTERS.
 CC -1- POLYMORPHISM: RHCE AND RHD ARE RESPONSIBLE FOR THE RH BLOOD GROUP
 CC SYSTEM. THE MOLECULAR BASIS OF THE E(UPPER CASE)-RH3/ELOWER
 CC CASE)-RH5 BLOOD GROUP ANTIGENS IS A SINGLE VARIATION IN POSITION
 CC 225. PRO-225 CORRESPONDS TO RH3 AND ALA-225 TO RH5. THE MOLECULAR
 CC BASIS OF THE C(UPPER CASE)-RH2/C(LOWER CASE)-RH4 BLOOD GROUP
 CC ANTIGENS IS A SINGLE VARIATION IN POSITION 102; SER-102
 CC CORRESPONDS TO RH2 AND PRO-102 TO RH4.
 CC -1- SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHD.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: M34015; AAA36567.1; -
 CC EMBL: X54534; CAA38401.1; -
 CC EMBL: S57967; AAB26080.1; -
 CC EMBL: X63096; CAA44809.1; -
 CC EMBL: X63096; CAA44810.1; -
 CC EMBL: X63096; CAA44812.1; -
 CC EMBL: S70456; AAD14061.1; -
 CC PIR: S02085; S02085.
 CC PIR: S02087; S02087.
 CC PIR: A32509; A32509.
 CC PIR: A30405; A30405.
 CC PIR: S13060; S13060.
 CC MIM: 111690; -
 CC MIM: 111700; -
 CC InterPro: IPR001905; -
 CC InterPro: IPR002229; -
 CC Pfam: PF00909; Ammonium_transp; 1.
 CC PRINTS: PR00342; RHESUSRD.
 CC Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing;
 CC Polymorphism.
 KW INIT MET 0 0
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 43 63 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 124 144 POTENTIAL.

FT	TRANSMEM	171	191	POTENTIAL.
FT	TRANSMEM	202	222	POTENTIAL.
FT	TRANSMEM	237	257	POTENTIAL.
FT	TRANSMEM	264	284	POTENTIAL.
FT	TRANSMEM	286	306	POTENTIAL.
FT	TRANSMEM	330	350	POTENTIAL.
FT	TRANSMEM	357	377	POTENTIAL.
FT	VARSPPLIC	162	312	MISSING (IN ISOFORM RHV1).
FT	VARSPPLIC	163	267	MISSING (IN ISOFORM RHV1).
FT	VARSPPLIC	313	353	VCCNRVGHHTISVMSISFLGLGELITRYIVLLVHTWN -> DWLPGPPHMGTOLGHRDSSHVSPPRFAPKSNMEST SCG (IN ISOFORM RHV1). MISSING (IN ISOFORM RHV1). IGFOVLISGELSLAIVIALTSGLLTLNLTAKRAPHVA KYFDVDFMKFPLAVGF -> FAPKSNMESTSC (IN ISOFORM RHV1). W -> C. /FTID=VAR_006911. A -> T (IN C(X)/RH9 ANTIGEN). /FTID=VAR_006912. O -> R (IN C(M)/RH8 ANTIGEN). /FTID=VAR_006913. L -> I. /FTID=VAR_006914. N -> S. /FTID=VAR_006915. P -> S (IN C/RH2 ANTIGEN). /FTID=VAR_006916. P -> A (IN E/RH5 ANTIGEN). /FTID=VAR_006917. L -> V (IN VS ANTIGEN). /FTID=VAR_006918. C -> L (IN REF. 6). D -> G (IN REF. 4). G -> C (IN REF. 4). SQ SEQUENCE 416 AA; 45429 MW; DC7AFAY9DCDCCE1F CRC64;

Query Match 86.0%; Score 1828; DB 1; Length 416;
 Best Local Similarity 88.7%; Pred. No. 5,7e-126;
 Matches 369; Conservative 12; Mismatches 35; Indels 0; Gaps 0;

QY	2	SCKYPBSVGRCLPLCALTLLEAALLLFFFTHYDASLEDDKGLVASYQVGGDTVTMAIG	61
DB	1	SSKYPBSVGRCLPLMALTLLEAALLLFFFTHYDASLEDDKGLVASYQVGGDTVTMAIG	60
QY	62	LGFLTSFRHRSSVAFNLFMLALGVQWALLDGLFSQPPGKVVITLPSIRLATMSAM	121
DB	61	LGFLTSFRHRSSVAFNLFMLALGVQWALLDGLFSQPPGKVVITLPSIRLATMSAM	120
QY	122	SVLISVAVIGKVLADLVVAVLEVTDLGNLRVAINIFFTVDYHMMHITYFAAYFGL	181
DB	121	SVLISAGAVIGKVLADLVVAVLEVTALGLNRVAINIFFTVDYHMMHITYFAAYFGL	180
QY	182	TVANCLPKPLPEGTEDDQATIPSLSAMGLFLMNEFSPNSALLSPRIERNAVNT	241
DB	181	TVANCLPKPLPKGTEDDDQRTIPSLSAMGLFLMNEFSPNSALLSPRIERNAVNT	240
QY	242	YYAVAVSVYTAISGSSSLAHPOGKISKTYGSHAVLPREGVAVTSCHLPSPLVIGLVA	301
DB	241	YYALAVSVYTAISGSSSLAHPOGKISKTYGSHAVLPREGVAVTSCHLPSPLVAVLVA	300
QY	302	GLISVFGKTYPGCCNNVYLIPHSISGVNYSLLGLEIITYIVLVVDYVAGNGMIGF	361
DB	301	GLISGAKCLPVCNNVYLIGHISVMSISFLGLGELITYIVLVLVHTWANGMIGF	360
QY	362	QVLLSIGELSLAIVIALTSGLLTLNLTAKRAPHKAYFDVDFMKFPLAVGF	417
DB	361	QVLLSIGELSLAIVIALTSGLLTLNLTAKRAPHKAYFDVDFMKFPLAVGF	416

RESULT 6
 RHLF_PANTR

ID RHLF_PANTR STANDARD; PRT; 416 AA.
 AC Q28426;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE RH-LIKE PROTEIN IIF (RHESUS-LIKE PROTEIN IIF).
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 NCBI_TaxID=9598;
 RX MEDLINE=95085595; PubMed=7993375;
 RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
 RA Colin Y., Ruffie J.;
 RT "Molecular genetics of chimpanzee Rh-related genes: their
 RT relationship with the R-C-E-F blood group system, the chimpanzee
 RT counterpart of the human rhesus system.";
 RL Biochem. Genet. 32:201-221(1994).
 CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
 CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L37048; AAA65622.1; -
 DR InterPro: IPR001905; -
 DR Pfam: PF000909; Ammonium_transp. 1.
 DR PRINTS: PR00342; RHESUSRHD.
 KW Erythrocyte; Transmembrane.
 FT INIT_MET 0
 FT TRANSMEM 11 31
 FT TRANSMEM 43 63
 FT TRANSMEM 76 96
 FT TRANSMEM 124 144
 FT TRANSMEM 171 191
 FT TRANSMEM 202 222
 FT TRANSMEM 237 257
 FT TRANSMEM 264 284
 FT TRANSMEM 286 306
 FT TRANSMEM 330 350
 FT TRANSMEM 357 377
 FT TRANSMEM 416 AA; 45136 MM; 325ED916268B4F2F CRC64;
 SQ SEQUENCE

Query Match 85.0%; Score 1806; DB 1; Length 416;
 Best Local Similarity 87.5%; Pred. No. 2,3e-124;

Matches 364; Conservative 11; Mismatches 41; Indels 0; Gaps 0;

QY 2 SCRYPRSVGRCPLGALTEALILFEFTHYDASLEDDKGVASYOVGODLTVMAG 61
 DB 1 SSKYPRSVRCPLGALTEALILFEFTHYDASLEDDKGVASYOVGODLTVMAG 60
 QY 62 LGELTSSFRHSHSSVAFNLFMLALGVOMAILLDGFIISOPPSKRVITLFSIALTMSAL 121
 DB 61 FGEFLTSSFRHSHSSVAFNLFMLALGVOMAILLDGFIISOPPSKRVITLFSIALTMSAL 120
 QY 122 SVLISDVAVIGKYNLAQVYVAVLVVAVTGLGNLHMVSNINFTNGYHMMMHMIIYFAAAYFGL 181
 DB 121 SVLISAGAVIGYVNLVQVYVAVLVVAVTGLGNLHMVSNINFTNGYHMMMHMIIYFAAAYFGL 180
 QY 182 TVAMCLPKPLPVECTEDDDQFATIPSLSAMIGALFLMVFPSVSALLRSIERKNVAVNT 241
 DB 181 TVAMCLPKPLPVDKEDKDDQIATIPSLSAMIGALFLMVFPSVSALLRSIERKNVAVNT 240

QY 242 YVAVAVSVTAISGSSSLAHPOGKISKTYGSHAVLPBGAVDTSCHLIPSEWLPVIGLVA 301
 DB 241 YVAAVSVTAISGSSSLAHPOGKISKTYGSHAVLPBGAVDTSCHLIPSEWLPVIGLVA 300
 QY 302 GLISVIGAKLPGCCNRVLCIPHSIMGVNFSLLGLLEIITVYLVLDVYVGNNGMIGF 361
 DB 301 GLISIGAKGPCCNVLCIPDSVHNFNFSLLGLLEIITVYLVLDVYVGNNGMIGF 360
 QY 362 QVLLSIGELSLAVIALTSLGALLNLKIRKAPAEKEDDOVEMKPFHLAVGF 417
 DB 361 QVLLSIGELSLAVIALTSLGALLNLKIRKAPAEKEDDOVEMKPFHLAVGF 416

RESULT 7

RHLC_GORGO

ID RHLF_PANTR STANDARD; PRT; 416 AA.

AC Q28426;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE RH-LIKE PROTEIN IC (RHESUS-LIKE PROTEIN IC).

OS Gorilla gorilla gorilla (Lowland gorilla).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.

NCBI_TaxID=9595;

RX MEDLINE=95085595; PubMed=7993375;

RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,

RA Colin Y., Ruffie J.;

RT "Molecular genetics of chimpanzee Rh-related genes: their

RT relationship with the R-C-E-F blood group system, the chimpanzee

RT counterpart of the human rhesus system.";

RL Biochem. Genet. 32:201-221(1994).

CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO

CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L37052; AAA65626.1; -

DR InterPro: IPR001905; -

DR Pfam: PF000909; Ammonium_transp. 1.

DR PRINTS: PR00342; RHESUSRHD.

KW Erythrocyte; Transmembrane.

FT INIT_MET 0

FT TRANSMEM 11 31

FT TRANSMEM 43 63

FT TRANSMEM 76 96

FT TRANSMEM 124 144

FT TRANSMEM 171 191

FT TRANSMEM 202 222

FT TRANSMEM 237 257

FT TRANSMEM 264 284

FT TRANSMEM 286 306

FT TRANSMEM 330 350

FT TRANSMEM 357 377

FT TRANSMEM 416 AA; 45129 MM; 0F3B623F9908E087 CRC64;

SQ SEQUENCE

Query Match 84.6%; Score 1798; DB 1; Length 416;

Best Local Similarity 86.5%; Pred. No. 8.7e-124;
 Matches 360; Conservative 15; Mismatches 41; Indels 0; Gaps 0;

```

QY 2 SCKYPRSVGRCPLCALTEALILFFFTHYDASLEDQGLVASYQVQODLTVAAG 61
  1 SSKYPRSVRCCLPLCALTEALILFFFTHYDASLEDQGLVASYQVQODLTVAAG 60
DB 62 LGFTSSFRHSSVAFNFLMALGVQMALDGLFLOPPSGKVITLESIRATSTM 121
  61 FGLTSSFRHSSVAFNFLMALGVQMALDGLFLOPPSGKVITLESIRATSTM 120
QY 122 SVLISDAVLGKYNLAOLVYVMEVTDLGNLRWVISINPTDYHMMMHIIYFAAYEGL 181
  121 SVLISDAVLGKYNLAOLVYVMEVTDLGNLRWVISINPTDYHMMMHIIYFAAYEGL 180
DB 182 TVAMCLPKPLPEGTEDNDQATIPSLSAMGALFLWMPFVSNSALLRSPDERKNAVENT 241
  181 TVAMCLPKPLPEGTEDNDQATIPSLSAMGALFLWMPFVSNSALLRSPDERKNAVENT 240
QY 242 YVAVAVSVTAISGSSLAHPQKISKTYGHSATVLPREGVAVDTSCHLIPSPWLPVLGLVA 301
  241 YVAVAVSVTAISGSSLAHPQKISKTYGHSATVLPREGVAVDTSCHLIPSPWLPVLGLVA 300
QY 302 GLISVFGAKLPGCCNRVLGIPHSISGYNFSLGLLEIIVYLVDVVGAGNGMIGF 361
  301 GLISVFGAKLPGCCNRVLGIPHSISGYNFSLGLLEIIVYLVDVVGAGNGMIGF 360
DB 362 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 417
  361 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 416
DB 361 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 416

RESULT 8
RHL_MACMU STANDARD; PRT; 416 AA.
AC 02849;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
  Cercopithecinae; Macaca.
NCBI_TaxID=9544;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=bone marrow;
RC MEDLINE=94223693; PubMed=7909570;
RA Moutou R., Le van Kim C., Cherif-Zahar B., Salvignol I., Blancher A.,
  Carlton J.-P., Colin Y.;
RT "Molecular characterization of the Rh-like locus and gene transcripts
  from the rhesus monkey (Macaca mulatta).";
RT J. Mol. Evol. 38:169-176(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
  HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation-
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb.ch/announce/
  CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S70343; AAB30637.1; -
DR InterPro: IPR001905; -
DR Pfam: PF00909; Ammonium transp. 1.
DR PRINTS: PR00342; RHESUSRD.
KW Erythrocyte; Transmembrane; Polymorphism.
FT INIT_MET 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.

```

```

FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT VARIANT 1 55 MISSING (IN MAC-B).
FT VARIANT 70 70 R -> K (IN MAC-B).
FT VARIANT 106 106 A -> V (IN MAC-B).
FT VARIANT 191 191 P -> L (IN MAC-B).
FT VARIANT 261 261 G -> R (IN MAC-B).
FT VARIANT 342 342 Y -> C (IN MAC-B).
SQ SEQUENCE 416 AA; 45672 MW; 3BFD939BEA79A6F CRC64;

Query Match 71.5%; Score 1519; DB 1; Length 416;
Best Local Similarity 73.8%; Pred. No. 1.7e-103;
Matches 307; Conservative 36; Mismatches 73; Indels 0; Gaps 0;

QY 2 SCKYPRSVGRCPLCALTEALILFFFTHYDASLEDQGLVASYQVQODLTVAAG 61
  1 SSKYPRSVRCCLPLCALTEALILFFFTHYDASLEDQGLVASYQVQODLTVAAG 60
DB 62 LGFTSSFRHSSVAFNFLMALGVQMALDGLFLOPPSGKVITLESIRATSTM 121
  61 FGLTSSFRHSSVAFNFLMALGVQMALDGLFLOPPSGKVITLESIRATSTM 120
QY 122 SVLISDAVLGKYNLAOLVYVMEVTDLGNLRWVISINPTDYHMMMHIIYFAAYEGL 181
  121 SVLISDAVLGKYNLAOLVYVMEVTDLGNLRWVISINPTDYHMMMHIIYFAAYEGL 180
DB 182 TVAMCLPKPLPEGTEDNDQATIPSLSAMGALFLWMPFVSNSALLRSPDERKNAVENT 241
  181 TVAMCLPKPLPEGTEDNDQATIPSLSAMGALFLWMPFVSNSALLRSPDERKNAVENT 240
QY 242 YVAVAVSVTAISGSSLAHPQKISKTYGHSATVLPREGVAVDTSCHLIPSPWLPVLGLVA 301
  241 YVAVAVSVTAISGSSLAHPQKISKTYGHSATVLPREGVAVDTSCHLIPSPWLPVLGLVA 300
QY 302 GLISVFGAKLPGCCNRVLGIPHSISGYNFSLGLLEIIVYLVDVVGAGNGMIGF 361
  301 GLISVFGAKLPGCCNRVLGIPHSISGYNFSLGLLEIIVYLVDVVGAGNGMIGF 360
DB 362 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 417
  361 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 416
DB 361 QVLISIGELSLAIVALTSGLTALLNLKIRKAPHEAKYFDQVFWKPHLAVGF 416

RESULT 9
RHL_MACFA STANDARD; PRT; 416 AA.
AC 02848;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
  Cercopithecinae; Macaca.
NCBI_TaxID=9541;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=bone marrow;
RC MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
  Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
  RT relationship with the R-C-E-F blood group system, the chimpanzee
  RT counterpart of the human rhesus system.";

```

```

RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L37054; AAA65628.1; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRD.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 45877 MW; 17822A558D2299A9 CRC64;

Query Match 70.88; Score 1504; DB 1; Length 416;
Best Local Similarity 74.38; Pred. No. 2e-102;
Matches 309; Conservative 31; Mismatches 76; Indels 0; Gaps 0;

QY 2 SKYPRSVGRCLPLCALTEALILLFYFTHDASLEDDKGLVASYOGODLTWMAIG 61
DB 1 SKYPRSVGRCLPLMALTEALILLFYFTHDASLEDDKGLVASYOGODLTWMAVLG 60
QY 62 LGFLTSSFRHRHSSVAFNLMALGVOMAILLDGFLSOPSPSKVVTITSLATMSAL 121
DB 61 LGFTSNLRHNSSVAFNLMALGVOMAILLDGFLSOPSPSKVYIKLSILATMSST 120
QY 122 SVLISDAVGLKYNLAOLVYVMEVTDGCLNLRMVISNIFNTDYHNMNMHIYFAAYEGL 181
DB 121 SMLISMAVGLKYNLAOLVYVMEVTEFVGTMRIVYINIFKIDYGNMNMHIHFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDNDQFATIPSLSAMGALFLMMPFSPVSALLRSPIERKNVFN 241
DB 181 TVAMCLPKPLPEGETEDKYQTTPSLFAMLGTLFLMFWFENFENSLINIERKNVFN 240
QY 242 YVAVAVSVAITSGSSLAHPQKISKTYGHSAYLPBGVAVDTSCHLIPSPWLPYVLGLVA 301
DB 241 YVALAVSAVTAIVSSSLAHQKINMTYMNAGLAGVAAGASCHVYHSHWIMAVGLVA 300
QY 302 GLISVGAATLPPCCNRYLCIPIHSSIMGVNFSLLGLEIITYLVLLDTVGAGNGMIGF 361
DB 301 GLISFGAKCLPYCFNVLGIHSHSMHYTFEGPALGETIYIVLMALRFVMASSNMIGF 360
QY 362 QVLTISGELAVIVALTSGLLTALLINKIKRAPHAKKFEVDQVFKPHLAVGF 417
DB 361 QVLTISGELSLAMAMSTISGLLTGLLNLKIMGPHAKYFDDQAEWEPHLAVGF 416

RESULT 10
RHL_HYLP1 STANDARD; PRT; 353 AA.
AC Q28446;
DT 01-NOV-1997 (Rel. 35, Created)

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
OS Hylobates pileatus (Pileated gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
ON NCBI_TaxID=9589;
RX [1]
RC TISSUE=bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socia W.W.,
RA Collin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RT Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L37051; AAA65625.1; -
DR InterPro: IPR001905; -
DR Pfam: PF00909; Ammonium_transp; 1.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0 BY SIMILARITY.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
SQ SEQUENCE 353 AA; 38417 MW; 3CFD25B38033AB6A CRC64;

Query Match 60.78; Score 1289; DB 1; Length 353;
Best Local Similarity 82.78; Pred. No. 7.2e-87;
Matches 258; Conservative 20; Mismatches 34; Indels 0; Gaps 0;

QY 2 SKYPRSVGRCLPLCALTEALILLFYFTHDASLEDDKGLVASYOGODLTWMAIG 61
DB 1 SKYPRSVGRCLPLMALTEALILLFYFTHDASLEDDKGLVASYOGODLTWMAIG 60
QY 62 LGFLTSSFRHRHSSVAFNLMALGVOMAILLDGFLSOPSPSKVVTITSLATMSAL 121
DB 61 LGFTSNLRHNSSVAFNLMALGVOMAILLDGFLSOPSPSKVYIKLSILATMSAM 120
QY 122 SVLISDAVGLKYNLAOLVYVMEVTDGCLNLRMVISNIFNTDYHNMNMHIYFAAYEGL 181
DB 121 SGLISGAVGLKYNLAOLVYVMEVTEFVGTMRVYININIFKIDYHNMNMHIHFAAYEGL 180
QY 182 TVAMCLPKPLPEGETEDNDQFATIPSLSAMGALFLMMPFSPVSALLRSPIERKNVFN 241
DB 181 TVAMCLPKPLPEGETEDKEQIANSPLSAMGALFLMFWFENFENSLINIERKNVFN 240
QY 242 YVAVAVSVAITSGSSLAHPQKISKTYGHSAYLPBGVAVDTSCHLIPSPWLPYVLGLVA 301
DB 241 YVALAVSAVTAIVSSSLAHQKINMTYMNAGLAGVAAGVAGTSCHLISSPWLMVGLVA 300
QY 302 GLISVGAATLPP 313
DB 301 GLISIGAKCLP 312

```


OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88273145; PubMed=2839489;
 RA Karpel R., Olami Y., Taglicht D., Schuldiner S., Padan E.;
 RT "Sequencing of the gene ant which affects the Na⁺/H⁺ antiporter
 activity in Escherichia coli.";
 RL J. Biol. Chem. 263:10408-10414(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1233-1238(1997).
 RN [4]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=92042080; PubMed=1657980;
 RA Karpel R., Alon T., Glasner G., Schuldiner S., Padan E.;
 RT "Expression of a sodium proton antiporter (NhaA) in Escherichia coli
 is induced by Na⁺ and Li⁺ ions.";
 RL J. Biol. Chem. 266:21753-21759(1991).
 RN [5]
 RP FUNCTION, AND PARTIAL SEQUENCE.
 RX MEDLINE=91250446; PubMed=1645730;
 RA Taglicht D., Padan E., Schuldiner S.;
 RT "Overproduction and purification of a functional Na⁺/H⁺ antiporter
 coded by nhaA (ant) from Escherichia coli.";
 RL J. Biol. Chem. 266:11289-11294(1991).
 RN [6]
 RP MUTAGENESIS OF HISTIDINES.
 RX MEDLINE=93165668; PubMed=8381959;
 RA Geerchman Y., Olami Y., Rimmon A., Taglicht D., Schuldiner S., Padan E.;
 RT "Histidine-226 is part of the pH sensor of NhaA, a Na⁺/H⁺ antiporter
 in Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1212-1216(1993).
 CC -1- FUNCTION: THIS IS A Na⁺/H⁺ ANTI-PORTER. IT CAN MEDIATE SODIUM
 UPTAKE WHEN A TRANSMEMBRANE PH GRADIENT IS APPLIED. IT PROBABLY
 ACTS IN THE REGULATION OF INTERNAL PH AT THE ALKALINE PH RANGE. IT
 CATALYZES THE EXCHANGE OF 2H⁺ PER NA⁺. ITS ACTIVITY IS HIGHLY
 DEPENDENT ON THE PH.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- INDUCTION: TRANSCRIPTION STIMULATED BY HIGH NA⁺ CONCENTRATIONS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a separate agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03879; AAA23448.1; ALT_INIT.
 CC EMBL: D10483; BAA01296.1; ALT_INIT.
 CC EMBL: AEO00112; AAC73130.1; -.
 CC EMBL: S67239; AAB20348.1; -.
 CC EMBL: A28800; A28800.
 CC PIR: S40541; S40541.

DR EcoGene; EGI0652; nhaA.
 KW Transmembrane; Inner membrane; Transport; Sodium transport; Symport.
 FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 205 221 POTENTIAL.
 FT TRANSMEM 254 274 POTENTIAL.
 FT TRANSMEM 283 299 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT SITE 225 225 IMPORTANT FOR PH SENSOR.
 FT CONFLICT 56 57 NK -> KQ (IN REF. 1).
 SQ SEQUENCE 388 AA; 41355 MW; B508BD2E5EE9130 CRC64;
 Query Match 5.6%; Score 119.5; DB 1; Length 388;
 Best Local Similarity 20.0%; Pred. No. 0.081;
 Matches 82; Conservative 69; Mismatches 143; Indels 117; Gaps 16;
 QY 18 LLEALILFFPTHTDASLEQKGLVASYQYGDITVMAAG-----LGFLLTSF-- 69
 DB 61 LWINDALMAVFFLVGLVEKRELMQGSLSLRQA--FPVIAAIGMIVPALVLAENYAD 119
 QY 70 --RRHSW-----SSVAFNLEMLA-LGVQMAILLDFLSQFP-----SGKVITLTF--SI 113
 DB 120 PIRREGAIPRAADIDIANALGVALLLSRSVPLAKIFLMALAIIDDGAILIILFTNDL 179
 QY 114 WLATMSALSVLISVDVAVLGKVNLAQLVVMVLEVTDLGNLRMVISNIENFTDYHMMNHIT 173
 DB 180 SMASLGVAAVAIAVLAVALNLCGARRTGVYILGVY-----VIMTAVLKSQGVH----- 225
 QY 174 VPAAYEGSLVAMOLKPKLPREGTEDNDQRTIIPSLSAMGLFLMFPSPNSNLSLSPTE 233
 DB 226 --ATLGVIVGFPTPLKEKKG-----NSPAK 249
 QY 234 RKNVENVTYAVAVSVYVTAISGSSLAHPQCKISKTYGSAVLPDEGVAVDTSCHLIPSPWL 293
 DB 250 RLEVLHPWVAVYLLPLFA-----FANAGVSQGVYLLDGLTSLP---- 289
 QY 294 PVLGVAGLISVYRKAYLPDCCNRVLGPRHSSIMGYNFSLLGLLEIYIVLLVLDTV- 352
 DB 290 ---LGIITAGLL-----IGKPLGISLPCWLAIRLKLAIHPGTYQQQIMVVGIIIC 335
 QY 353 GAGNGMIGPVLLSIGELSLAIY-----IALTSGILRAILLN---IKIRKAP 396
 DB 336 GIGFTMSIFIASLAFGSDPEELINMAKLGILVGSISSAVIGYSWLRVRLRP 386
 RESULT 13
 ID YJ78_YEAST STANDARD; PRT; 881 AA.
 AC P39535;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOHETICAL 97.7 KDA MEMBRANE PROTEIN IN PRP21-UBP12 INTERGENIC
 DE REGION.
 GN YUL198W OR J03336.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE=95274326; PubMed=7754713;
 RA Purnelle B., Coster F., Goffeau A.;
 RT "The sequence of a 36 kb segment on the left arm of yeast chromosome
 X identifies 24 open reading frames including NUC1, PRP21 (SPP31),
 RT CDC6, CR2, the gene for S24, a homologue to the acetylase gene ACO1
 and two homologues to chromosome III genes.";

RL Yeast 10:1235-1249(1994).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE NAD(P)/PHO87 FAMILY OF TRANSPORTERS.
 CC PHO87 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X77688; CAA54759.1; -
 DR EMBL: 249473; CAA89493.1; -
 DR PIR: S46633; S46633.
 DR SGD: S0003734; YJUL198W.
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 FT TRANSMEM 493 513 POTENTIAL.
 FT TRANSMEM 514 534 POTENTIAL.
 FT TRANSMEM 539 559 POTENTIAL.
 FT TRANSMEM 581 601 POTENTIAL.
 FT TRANSMEM 663 683 POTENTIAL.
 FT TRANSMEM 691 711 POTENTIAL.
 FT TRANSMEM 718 738 POTENTIAL.
 FT TRANSMEM 758 778 POTENTIAL.
 FT TRANSMEM 805 825 POTENTIAL.
 FT TRANSMEM 854 874 POTENTIAL.
 SQ SEQUENCE 881 AA; 97688 MW; BB4466BCAD176714 CRC64;

Query Match 5.6%; Score 119; DB 1; Length 881;
 Best Local Similarity 20.5%; Pred. No. 0.21; Mismatches 130; Indels 116; Gaps 17;

Matches 80; Conservative 65; Mismatches 130; Indels 116; Gaps 17;
 QY 86 LGVOMA-----ILDGF-----LSQFPGSKV-----VITLFSIW 114
 DB 494 LAAMWSSTIMILLAGFTLGEVLAQYNAKVLASWLAFAQCKPRNVLLMAMCVFFFLSMW 553
 QY 115 LA-----TWSALSVLSIDAVLGKYNLAQVVMVLEVTdGCLNLRVVISNIFNTDYM 167
 DB 554 ISNVAALVLYSTLSPL--DAMDADSPFAQALVGLAALANIGMSSPISPNONI--I 608
 QY 168 NMHIYFAAYFGLTVMCLPKPLPEGTEDDDQATIPSLAMGALEFMW-----PEPS 222
 DB 609 SMSYLKPYGIGWGQFFVALP-----SGILLMLLWILLFTTERKM 649
 QY 223 VNSALRSPIERKNAVENTYAAVAVSVTAISGSLAHPOGKISKTYGHSALP-----EG 278
 DB 650 KTKLEKRPKIKTKTY-KQYIIIVYATILLMCVEQIEBAPSS-QQIAIIRIVLFFG 707
 QY 279 VAVDTSCHLIPSPWLPIVLG-----VAGLISVYGAQYLP-----CCN 317
 DB 708 TGLSTODLNAFPIVILAMGIALGKAVSSGLSTI-AKALQKRIENDGVAIICIF 766
 QY 318 RVLGIPHSSTMGVNF--LLGLLE-----IIVYLIVDTVAGAGMIGFO 362
 DB 767 GILMLVGVTFVSHVSAIIIIPIVQEVGDKLGNPKAAPILVEGCALLSSCGMGLASSGFP 826
 QY 363 VLSTI-----GELSLAVIALTSGLTALL 387
 DB 827 NVTALSKVDRKGRYLSVMTFLTRGVPASIL 857
 RESULT 14
 ID Y092_HAEIN STANDARD; PRT; 419 AA.
 AC 057493;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HYPOTHETICAL PROTEIN HI0092.
 GN HI0092.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=9530630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kierulff A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Ufferback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.U., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd.";
 RT Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32694; AAC2170.1; -
 DR TIGR: HI0092; -
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 24 44 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 101 121 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 216 236 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 SQ SEQUENCE 419 AA; 42781 MW; 6DCA54BB0A55FE73 CRC64;

Query Match 5.3%; Score 112.5; DB 1; Length 419;
 Best Local Similarity 19.8%; Pred. No. 0.28; Mismatches 145; Indels 113; Gaps 17;

Matches 81; Conservative 71; Mismatches 145; Indels 113; Gaps 17;

QY 43 GLVASVYGDQDLYMAAIGLGLTSFRRHSWSSVAFNLMALGVOMAILD-----GFL 98
 DB 38 GLIGCALDSQVTSIMIGAGQITTAVMK-----ILAGVLAGVLDIESGAANSI 85
 QY 99 SQFSGKVVIT--LESLATMSALSVLSYDAVGLKYNLAQVVMVLEVTdGCLNLRV 156
 DB 86 TETITNKLGETRALLALALAMILLTAAGVFDVAIVTS--PALALSRSDLSKAIL 142
 QY 157 IS-----NIFNTD-----YHM--NMHIYFAAYFGLVAMCLPKPL-PEGT 195
 DB 143 LAMIGGKAGNIMSPNPAIAADTFHPLPLSVMMAGIIPALFOLLILYFLAKLINKGS 202
 QY 196 EDNDQFA-----TIPS-----LSAMGALFLMFPSPVNSALRSPIERKNAVENTYAAV 245
 DB 203 KVTDEKIVILETQMLPSPFLVLPVLAIVLILALRP-----LFDI 242
 QY 246 AVSVVTAISGSSL--AHPOGKISKTYGHSALVPGVAVDTSCHLIPSPWLPIVLGVL 303
 DB 243 KVDPLIALPLGLGLIGAGFCMGKLRINISY-----AINGLSKMTPVAIMLIGTALAGI 294

OY		ISVIAKVLPGCCNRNVLGIIPHSIMGVENSLGLLEIIVLYLVLDTRVGAGMNGIQOY	363
Dd	295	IANSLSKEY-----LIQIGHSGDSTSY-----LLAPISG-----V	324
Oy	364	LIISELGSILAIVALTSGLTALTALLNLTKRPAHEAK-YFDQVEFWKEPPH	412
Dd	325	LMSLATPASTTAGTAVASNVFSSLTLELGGVSSLAGAAMIHACATVPVDHNPH	374
RESULT	15		
ID	GNTT_ECOLI	STANDARD:	PRT; 437 AA.
AC	P39835;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	HIGH-AFFINITY GLUCONATE TRANSPORTER (GLUCONATE PERMEASE) (GNT-I SYSTEM).		
GN	GNTT OR USGA OR GNTM.		
OS	Escherichia coli.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OX	NCBI_TaxID=562;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12 / MG1655;		
RX	MEDLINE=97426617; PubMed=9278503;		
RA	Blatter F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;		
RL	"The complete genome sequence of Escherichia coli K-12."		
RM	Science 277:1453-1474(1997). [2]		
RN	PRELIMINARY SEQUENCE OF 381-437 FROM N.A.		
RP	MEDLINE=89013888; PubMed=2845225;		
RX	Pugsley A.P., Dubreuil C.;		
RA	"Molecular characterization of malQ, the structural gene for the Escherichia coli enzyme amyloamylase.";		
RL	Mol. Microbiol. 2:473-479(1988). [3]		
RN	IDENTIFICATION.		
RP	MEDLINE=95075659; PubMed=7984428;		
RX	Borodovsky M., Rudd K.E., Kooin E.V.;		
RA	"Intrinsic and extrinsic approaches for detecting genes in a bacterial genome.";		
RL	Nucleic Acids Res. 22:4756-4767(1994). [4]		
RN	CHARACTERIZATION.		
RP	MEDLINE=87310381; PubMed=3040894;		
RX	Isturiz T., Palmero E., Vitelli-Flores J.;		
RA	"Mutations affecting gluconate catabolism in Escherichia coli.		
RL	Genetic mapping of the locus for the thermosensitive gluconokinase."; J. Gen. Microbiol. 132:3209-3219(1986). [5]		
RN	CHARACTERIZATION.		
RP	MEDLINE=97197521; PubMed=9045817;		
RA	Porco A., Peekush N., Bausch C., Tong S., Isturiz T., Conway T.;		
RL	"Molecular genetic characterization of the Escherichia coli gntJ gene of GntC, the main system for gluconate metabolism.";		
J. Bacteriol. 179:1584-1590(1997).	-1 FUNCTION: PART OF THE GLUCONATE UTILIZATION SYSTEM GNT-I; HIGH-		
AFFINITY INTAKE OF GLUCONATE.	-1 PATHWAY: GLUCONATE UTILIZATION.		
-1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE (POTENTIAL).	-1 SIMILARITY: BELONGS TO THE GNP FAMILY OF PERMEASES.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial		

```

CC   entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL; U18997; AAN58213.1; -.
DR   EMBL; AE000417; AAC76440.1; -.
DR   EMBL; M32793; -. NOT_ANNOTATED_CDS.
DR   EcoGene; EG12380; gntT.
KM   Glucanate utilization; Sugar transport; Transport; Transmembrane;
KM   Inner membrane; Multigene family.
FT   TRANSMEM      2      22      POTENTIAL.
FT   TRANSMEM      23      43      POTENTIAL.
FT   TRANSMEM      50      70      POTENTIAL.
FT   TRANSMEM      107     127     POTENTIAL.
FT   TRANSMEM      133     153     POTENTIAL.
FT   TRANSMEM      173     193     POTENTIAL.
FT   TRANSMEM      221     241     POTENTIAL.
FT   TRANSMEM      257     277     POTENTIAL.
FT   TRANSMEM      291     311     POTENTIAL.
FT   TRANSMEM      326     346     POTENTIAL.
FT   TRANSMEM      348     368     POTENTIAL.
FT   TRANSMEM      369     389     POTENTIAL.
FT   TRANSMEM      417     437     POTENTIAL.
FT   CONFLICT      384     384     S -> A (IN REF. 2).
FT   CONFLICT      419     420     ET -> VS (IN REF. 2).
SQ   SEQUENCE      437 AA;  45923 MW;  2AE6F8BD049D66 C6C64;

```

Query Match	5.28;	Score 111.5;	DB 1;	Length 437;
Best Local Similarity	20.58;	Pred. No. 0.35;		
Matches	93;	Mismatches 152;	Indels 141;	Gaps 20;
		Conservative 68;		

```

01 56 VMAALGJOLFSSRRRHSWS--VAFNEFMJALQVOMILLDGLSOPSGKVITLFSW 114
02 4 VIVAIGVILLLLLLMIRRMNGFIALVLAVALGMSGMRPDKYIGSKAG----- 53
03 115 LATMSALSVLISVDVALGKY---NLAOLVVMVLVEVTDIGNR----- 154
04 54 VADVGSIALINGFAMGKMLADGSGAOTIATTLIAFGKKNHOMAVLITGFVSGFALY 113
05 155 ---MYSINFEINDYHNMNMHITY---FAAYGELVAKCLRPRLRGSTDDNOQTIES 206
06 114 EVGEVLMLPRVFTTAAASNPRLDLVGVPMAAALSVTHGFLRPHRGPAI---ATI-- 165
07 207 L$AMIGALFEMFPR$VSYNALSRLPIERKNAVENTYVAVSVTAISGSSLAHPGKIS 266
08 166 FNAODGKLLIXGLTILAIPTVLIAPV-----YA-----RYLKGIDRKIRPGLXS 209
09 267 -KTYGHSNVLPREGVAVDTHSCHLIR$PMLRPIVL--GLVAGLISVFGAKYIPGCCNRVLGI 322
10 210 AKTESSEEMPRFSGSVYVTSL-----VPVLLMAMRAIAEMIRPKGHAFLP--VAEFLTD 260
11 323 PHSSIMGCVN$SL-----GL-----L$EIIYVLLVLPDVGNGNMGIFOYL 364
12 261 P---VMATLIVLAIAMFTTFLGNRGSRMDQINDTLVSSIKIIMAMLLIIGGGAF--KOYL 315
13 365 L$IG-----EL$LAIVIALTISGLTALLAN----- 389
14 316 VDSGVDKYIASMHETNISPLLMAM$IAVLRITALGSATVAALTAGIAPLIATTVG$P 375
15 390 -----LKITKAPHEAKFYDDOVFKMFP---HLANG 416
16 376 ELMAIVAGSGSVIF$HVNDDPFWLFEKYE$FNLITJ 409

```

Search completed: September 12, 2001, 07:42:42
Job time: 211 sec

THIS PAGE BLANK (USPTO)


```

Db 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
Qy 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
|:|||||
Db 181 LSVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Db 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
Qy 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
|:|||||
Db 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
Qy 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417
|:|||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417

RESULT 2
ID 09U021 PRELIMINARY: PRT: 417 AA.
AC 09U021:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHLYO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TO;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDya in Japanese.";
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB018968; BAA81501.1;
DR InterPro: IPR001905;
DR InterPro: IPR002229;
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45177 MW; B72337B16747E6ID CRC64;

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 94.7%; Pred. No. 6.3e-118;
Matches 395; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MSCYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
|:|||||
Db 1 MSSYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Qy 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
Qy 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
|:|||||
Db 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
Qy 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
|:|||||
Db 181 LSVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Db 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
Qy 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
|:|||||
Db 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360

```

```

Db 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
Qy 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417
|:|||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417

RESULT 3
ID 09U020 PRELIMINARY: PRT: 417 AA.
AC 09U020:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHLYO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TO;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDya in Japanese.";
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB018969; BAA82159.1;
DR InterPro: IPR001905;
DR InterPro: IPR002229;
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45179 MW; 62B11DD24064F7D5 CRC64;

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 95.0%; Pred. No. 6.3e-118;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MSCYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
|:|||||
Db 1 MSSYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Qy 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
Qy 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
|:|||||
Db 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
Qy 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
|:|||||
Db 181 LSVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Qy 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Db 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
Qy 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
|:|||||
Db 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
Qy 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417
|:|||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417

```

```

RESULT 4
ID 09UK19 PRELIMINARY: PRT: 417 AA.
AC 09UK19:
DT 01-MAY-2000 (TREMBlrel. 13, Created)

Query Match 93.0%; Score 1976; DB 4; Length 417;
Best Local Similarity 95.0%; Pred. No. 6.3e-118;
Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy 1 MSCYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
|:|||||
Db 1 MSSYPRSVRCPLCLALTEALILILEFFTHYDASLEDDKGLVASYOGODLTWMAI 60
61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Qy 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
|:|||||
Db 61 GLGFLTSSFRHSSWSVAFLMFLALGVQWAILDGLFSQPSGKVYITLFSRLATMSA 120
Qy 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
|:|||||
Db 121 LSVLISYDAVLGKYNLAQLVVWVLEVTALGNLRMVISNIFNTDYHMMMHIIYVFAAYFG 180
Qy 181 LTVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
|:|||||
Db 181 LSVAMCPLPKPLPECTEDDDQATIPSLSAMGLFLMFPVSVALRSPIERKNVFN 240
241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Qy 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
|:|||||
Db 241 TYVAVAVSVTVAISGSSLAHPQKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
Qy 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
|:|||||
Db 301 AGLISVGAATLPPCCNRCVIGIPHSSTMGYNFSLGLGELIYIVLLVDTVGAGNGMIG 360
Qy 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417
|:|||||
Db 361 FOVLSTIGELSLAIVIALTSGLLTALLNKIRKAPHEAKYFDDQVFWKPHLAVGF 417

```

DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DE 01-JUN-2000 (Tremblrel. 14, last annotation update)
 DE RHD TYPE IITA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97398395; PubMed=9256293;
 RA Huang C.H., Chen Y., Reid M.;
 RT "human D(IIa) erythrocytes: Rhd protein is associated with multiple
 RT dispersed amino acid variations.";
 RL Am. J. Hematol. 55:139-145(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Huang C.H.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF187846; AAF04565.1; -;
 DR InterPro; IPR001905; -;
 DR Pfam; PF00909; Ammonium_transp; 1.
 DR PRINTS; PR00342; RHESUSRHD.
 SQ SEQUENCE 417 AA; 45273 MW; 491FB1FE1BA37957 CRC64;

 Query Match 92.9%; Score 1975; DB 4; Length 417;
 Best Local Similarity 95.0%; Pred. No. 7.3e-118;
 Matches 396; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

 QY 1 MSCKYPSVGRCLPLCALTEALILFFFTHYDASLEDOKGLVASVOGDLTVMAAI 60
 Db 1 MSKYPSVGRCLPLMALTEALILFFFTHYDASLEDOKGLVASVOGDLTVMAAL 60

 QY 61 GLGFLTSSFRHSSVAFLNFMALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120
 Db 61 GLGFLTSSFRHSSVAFLNFMALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120

 QY 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
 Db 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

 QY 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
 Db 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

 QY 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240
 Db 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240

 QY 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300
 Db 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300

 QY 301 AGLISV-GAKYLPCCCNRVLGIPHSIMGVNFSILGLLEITIVLVLDVVGANGMIG 360
 Db 301 AGLISVGGAKYLPCCCNRVLGIPHSIMGVNFSILGLLEITIVLVLDVVGANGMIG 360

 QY 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWKPFPHLAVGF 417
 Db 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWKPFPHLAVGF 417

 RESULT 5
 Q9U022 PRELIMINARY; PRT; 417 AA.
 AC Q9U022;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
 DE RH BLOOD GROUP D ANTIGEN (RHD).
 GN RHDA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=TO;
 RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
 RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
 RT "Polymorphisms of Rhdya in Japanese.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB018967; BA681900.1; -;
 DR InterPro; IPR001905; -;
 DR InterPro; IPR002229; -;
 DR Pfam; PF00909; Ammonium_transp; 1.
 DR PRINTS; PR00342; RHESUSRHD.
 SQ SEQUENCE 417 AA; 45211 MW; 764B092CBF95ECC4 CRC64;

 Query Match 92.9%; Score 1974; DB 4; Length 417;
 Best Local Similarity 94.7%; Pred. No. 8.5e-118;
 Matches 395; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

 QY 1 MSCKYPSVGRCLPLCALTEALILFFFTHYDASLEDOKGLVASVOGDLTVMAAI 60
 Db 1 MSKYPSVGRCLPLMALTEALILFFFTHYDASLEDOKGLVASVOGDLTVMAAI 60

 QY 61 GLGFLTSSFRHSSVAFLNFMALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120
 Db 61 GLGFLTSSFRHSSVAFLNFMALGVOMAILLDGFLSQPPSGKVITLFSRLATMSA 120

 QY 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180
 Db 121 LSVLISVDAYGKYNLAQLVVMVLEVTGLGNLBMVISNIFNFDYHNMHHIYFAAYFG 180

 QY 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240
 Db 181 LTVAMCLPKPLPECTEDNDQFATIPSLAMGALFLMFPSPVSALLRPIERKNVFN 240

 QY 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300
 Db 241 TYAVAVSVTAISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIPSPWLPVIGLV 300

 QY 301 AGLISV-GAKYLPCCCNRVLGIPHSIMGVNFSILGLLEITIVLVLDVVGANGMIG 360
 Db 301 AGLISVGGAKYLPCCCNRVLGIPHSIMGVNFSILGLLEITIVLVLDVVGANGMIG 360

 QY 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWKPFPHLAVGF 417
 Db 361 FOVLISIGELSLAIVIALTSGLTALLNLKIRAPHAKEFFDQVWKPFPHLAVGF 417

 RESULT 6
 Q9NXY8 PRELIMINARY; PRT; 417 AA.
 AC Q9NXY8;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE RHESUS BLOOD GROUP, D ANTIGEN.
 GN RHD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wagner F.F., Flegel W.A.;
 RT "Molecular structure of weak D type 18, another RHD allele detected
 RT among weak D phenotypes.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ287289; CAB75731.1; -;
 DR EMBL; AJ287290; CAB75731.1; JOINED.
 DR EMBL; AJ287291; CAB75731.1; JOINED.
 DR EMBL; AJ287292; CAB75731.1; JOINED.
 DR EMBL; AJ287293; CAB75731.1; JOINED.
 DR EMBL; AJ287294; CAB75731.1; JOINED.
 DR EMBL; AJ287295; CAB75731.1; JOINED.

Db 181 LSVAMCLPKPLPECTEDKQDTATIPSLSAMGLFMMFSPSALLRSPERKNAFVN 240
 QY 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 Db 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 QY 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360
 Db 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360
 QY 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEHLAVGF 417
 Db 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEHLAVGF 417

RESULT 9
 Q9H248 PRELIMINARY; PRT; 417 AA.
 AC Q9H248;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE RHESUS D CATEGORY VI TYPE IV PROTEIN.
 GN RHD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Esteban R., Nogues N., Montero R., Hurtado M., Maroto S., Ribera A.;
 RT "Characterization of a new RHD category VI (type IV) allele."
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBD databases.
 DR EMBL: AF312679; AAC40872.1; -
 SQ SEQUENCE 417 AA; 45293 MW; 03FD058960E41D89 CRC64;

Query Match 91.4%; Score 1942; DB 4; Length 417;
 Best Local Similarity 93.0%; Pred. No. 9e-116;
 Matches 386; Conservative 8; Mismatches 21; Indels 0; Gaps 0;
 QY 1 MSCKYPSVSGKCLPLCALTLLEAAILLFFPTHYDASLEDDKGLVASVYQGDLTVMMAI 60
 Db 1 MSCKYPSVSGKCLPLCALTLLEAAILLFFPTHYDASLEDDKGLVASVYQGDLTVMMAI 60
 QY 61 GLGFLTSSFRHSHSSVAFNLFMLALGVOMAILLDGFLSQPPSGKVITLFSIRLATWSA 120
 Db 61 GLGFLTSSFRHSHSSVAFNLFMLALGVOMAILLDGFLSQPPSGKVITLFSIRLATWSA 120
 QY 121 LSVLISDAVLGKYNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 Db 121 LSVLISDAVLGKYNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 QY 121 MSVLISGAVLGNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 Db 121 MSVLISGAVLGNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 QY 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMWFPSVNSALLRSPERKNAFVN 240
 Db 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMWFPSVNSALLRSPERKNAFVN 240
 QY 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 Db 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 QY 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360
 Db 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360
 QY 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEHLAVGF 417
 Db 361 FOVLISIGELSLAIVIALTSGLLTALLNLKIRAPHEAKYFDDQVFWKPEHLAVGF 417

RESULT 10
 Q9UD25

ID Q9UD25 PRELIMINARY; PRT; 417 AA.
 AC Q9UD25;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE RHESUS D CATEGORY VI TYPE III PROTEIN.
 GN RHD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD.
 RX MEDLINE=98158656; Pubmed=9490704;
 RA Wagner F.F., Gassner C., Mueller T.H., Schoenitzer D., Schunter F.,
 RA Flegel W.A.;
 RT "Three molecular structures cause Rhesus D category VI phenotypes with
 RT distinct immunohematologic features."
 RL Blood 91:2157-2168(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD.
 RX MEDLINE=9425182; Pubmed=8188244;
 RA Cherif-Zahar B., Le Van Kim C., Rouillac C., Raynal V., Cartton J.,
 RA Colin Y.;
 RT "Organization of the gene (RHCE) encoding the human blood group RHCE
 RT antigens and characterization of the promoter region."
 RL Genomics 19:68-74(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD.
 RX MEDLINE=9306356; Pubmed=1438298;
 RA Le Van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
 RA Cartton J.P., Colin Y.;
 RT "Molecular cloning and primary structure of the human blood group RHD
 RT polypeptide."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
 DR EMBL: 297026; CAB09722.1; -
 DR InterPro: IPR001905; -
 DR InterPro: IPR002229; -
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PRINTS: PR00342; RHESUSRHD.
 SQ SEQUENCE 417 AA; 45247 MW; 1887A3C519149E8F CRC64;

Query Match 90.9%; Score 1932; DB 4; Length 417;
 Best Local Similarity 92.6%; Pred. No. 3.9e-115;
 Matches 386; Conservative 9; Mismatches 22; Indels 0; Gaps 0;
 QY 1 MSCKYPSVSGKCLPLCALTLLEAAILLFFPTHYDASLEDDKGLVASVYQGDLTVMMAI 60
 Db 1 MSCKYPSVSGKCLPLCALTLLEAAILLFFPTHYDASLEDDKGLVASVYQGDLTVMMAI 60
 QY 61 GLGFLTSSFRHSHSSVAFNLFMLALGVOMAILLDGFLSQPPSGKVITLFSIRLATWSA 120
 Db 61 GLGFLTSSFRHSHSSVAFNLFMLALGVOMAILLDGFLSQPPSGKVITLFSIRLATWSA 120
 QY 121 LSVLISDAVLGKYNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 Db 121 LSVLISDAVLGKYNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 QY 121 MSVLISGAVLGNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 Db 121 MSVLISGAVLGNLAQLVVMVLEVTGLGNLMMVINSIFNTDYHNMHMIYFAAYFG 180
 QY 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMWFPSVNSALLRSPERKNAFVN 240
 Db 181 LTVAMCLPKPLPECTEDDQATIPSLSAMGLFLMWFPSVNSALLRSPERKNAFVN 240
 QY 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 Db 241 TYVAVAASVTAATSGSSLAHPQCKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVGLV 300
 QY 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360
 Db 301 AGLSVAGAKYLPCCNRRVIGIPHSISMGYNFSLGLLGLLTIYVLLVDVVGANGMIG 360

Oy	361	FOVLLISIGELSLAIYALTSGLTALLNLKIRKAPHEKKYRDDOVFKMFPHLANGF	417
Dd	361	FOVLLISIGELSLAIYALTSGLTALLNLKIKWAKPHAKYRDDOVFKMFPHLANGF	417
 RESULT 11			
ID	Q9UK74	PRELIMINARY:	PRT: 417 AA.
AC	Q9UK74:		
Dt	01-MAY-2000 (TREMBLrel. 13, Created)		
Dt	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
Dt	01-MAR-2001 (TREMBLrel. 16, Last annotation update)		
De	RH BLOOD GROUP ANTIGEN EVANS.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCBI_TaxID=9606;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PERIPHERAL BLOOD;		
RX	MEDLINE=96404415; PubMed=8808597;		
RA	Huang C.H., Chen Y., Reid M., Ghosh S.;		
RT	"Genetic recombination at the human RH locus: a family study of the		
RT	red-cell Evans phenotypic reveals a transfer of exons 2-6 from the RHD		
RT	to the RHCE gene.";		
Am	J. Hum. Genet. 59:825-833(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PERIPHERAL BLOOD;		
RA	Huang C.H.;		
RL	Submitted (AUG-1999) to the EMBL/Genbank/DDBJ databases.		
Df	EMBL: AF177938; AAD55773.1; -		
DR	InterPro: IPR001762; -		
DR	InterPro: IPR001905; -		
DR	InterPro: IPR002229; -		
DR	Pfam: PF00909; Ammonium_transp. 1.		
DR	PRINTS: PR00342; RHESUSRHD.		
DR	Prodor: PD00664; -; 1.		
SQ	SEQUENCE 417 AA: 45445 MW: DBAAD148ADC73F7 CRC64:		
 Query Match 89.3%; Score 1897; DB 4; Length 417; Best Local Similarity 92.1%; Pred. NO. 6.4e-113; Matches 384; Conservative 5; Mismatches 28; Indels 0; Gaps 0			
Oy	1	MSCKPRPSVGRPLCATLTLEAILILFFPFHYDASLEDOKGLVASVOGODLYMAAI	60
Dd	1	MSKKPRSRRCRLPCATLTLEAILILFFFYHYDASLEDOKGLVASVOGODLYMAL	60
Oy	61	GIGFLTSSFRRHSWSVAFNLFMLAGVOMAILLDGFLSOPFSRGKVILFFSTLATMSA	120
Dd	61	GIGFLTSSFRRHSWSVAFNLFMLALGVOMAILLDGFLSQFPGSKVITTLFSTRLATMSA	120
Oy	121	LSVLISVDALGVKLAAQLVNVVLVEVTDLGNLRVVISINFMTDYHMNMHIYFAAYFG	180
Dd	121	LSVLISVDALGVKLAAQLVNVVLVEVTALGNLRVVISINFMTDYHMNMHIYFEAFYFG	180
Oy	181	LTVANCLRKPRPEGEEDNDQFTITSLSAMLGATFLMKPSPSVNSALLRSPIERKNAYVN	240
Dd	181	LSVANCLRKPRPEGEEDNDQFTITPSLSAMLAGLFLMKMPFSNSALLRSPIERKNAYVN	240
Oy	241	TYYAAAVSWTAIFSGSSLAHPOGKISKTYTHSAVLPREGAVASTSCHLIPSPMLPIYLGLV	300
Dd	241	TYYAAAVSWTALISSSLAHPOKRISKITYVHSAVLAVAGVAVGSTCHLIPSPMLANVGLV	300
Oy	301	AGLISIVGAKYLPGCCNRVLAGIPHSSIMGYNFSLGLLEIITYLVLLVDTVGAGNGMG	360
Dd	301	AGLISIVGAKYLPVCNCRVLIHNHSVMHSIFSLLGLLCIEITYLVLLVDHTWNGNGMG	360
Oy	361	FOVLLISIGELSLAIYALTSGLTALLNLKIRKAPHEKKYRDDOVFKMFPHLANGF	417
Dd	361	FOVLLISIGELSLAIYALTSGLTALLNLKIWKAPRHAKYRDDOVFKMFPHLANGF	417

[illegible]

RESULT	14		
09UEC3			
ID	09UEC3	PRELIMINARY;	PRT: 417 AA.
AC	09UEC3		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-MAR-2001 (Tremblrel. 16, Last annotation update)		
DE	RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.		
GN	RHCE.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_Taxid=9606;		
	[1]		
RN	SEQUENCE FROM N.A.		
RA	Kashiwase K., Ishikawa Y., Uchikawa M.;		
RT	"The serological profile and molecular basis of the RHe variants (RHEFM and RHEKH) in Japanese.";		

ID	Q9UEC2	PRELIMINARY;	PRT;	417 AA.
AC	Q9UEC2;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	RH BLOOD GROUP CE ANTIGEN POLYPEPTIDE.			
GN	RHCE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kashivase K., Ishikawa Y., Uchikawa M.;			
RT	"Rhe serological profile and molecular basis of the Rhe variants			
RL	(RHEFM and RHEKH) in Japanese."			
DR	Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; ABO18645; BAA33828.1; "			
DR	InterPro: IPRO01905; "			
DR	InterPro: IPRO02229; "			
DR	Pfam: PF00909; Ammonium_transp. 1.			
DR	PRINTS; PR00342; RHESUSRHD.			
SO	SEQUENCE 417 AA; 45505 MW; 67B53E74C8B053DE CRC64;			
Query Match	86.0%;	Score 1827;	DB 4;	Length 417;
Best Local Similarity	88.5%;	Pred. No. 1,8e-108;		

Matches 369; Conservative 12; Mismatches 36; Indels 0; Gaps 0;

```
QY 1 MSCTPRSVGRCLPLCALTEALILFEFTHDASLEDOKGLVASOVGODLTVAAI 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSSKTPRSVRCLPLMALTEALILFEFTHDASLEDOKGLVASOVGODLTVAAL 60
QY 61 GLGFLTSSFRHSHSVAFNLFMLALGVMAILLDGLSQFPSPGKVVITLFSIWLATMSA 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GLGFLTSMFRHSHSVAFNLFMLALGVMAILLDGLSQFPSPGKVVITLFSIRLATMSA 120
QY 121 LSVLISYDAVLGKYNLAQLVYVMVLEVTGLGNLFMVISNFTNDYHMMNMHIYFAAYFG 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 MSVLISAGAVLGKYNLAQLVYVMVLEVTALGTLTVMVISNFTNDYHNNLRHFYVFAAYFG 180
QY 181 LTVAMCLPKPLPECTEDNDQFATIPSLSAMGALFLMMP-PSVNSALLRSPIERKNAVFN 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LTVAMCLPKPLPKCTEDNDQFATIPSLSAMGALFLMMPFVSNSPLLRSPIQKKNMVFN 240
QY 241 TYVAVASVYTAISGSSLAHPQGRISKTYGHSAVLPQGVAVDTSCHLIPSPWLPVLGLV 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TYVALAVSVYTAISGSSLAHPQGRISMTYHSAVLGAGVAVGTCHLIPSPWLAIVGLV 300
QY 301 AGLISVCGAKYLPCCCNRVLCIPHSSTMGYNFSLGLLAEIYIVLLVDTVGAGNGMIG 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 AGLISIGGAKCLPYCCNRVLCIHHISYVHSIFSLLGLGETIYIVLLVHTVMNGMIG 360
QY 361 FOVLSTIGELSLAIVIALTSGLTALLNLKIKRAPHAKYFDDOVFMKPPHLAVGF 417
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 FOVLSTIGELSLAIVIALTSGLTALLNLKIKRAPHAKYFDDOVFMKPPHLAVGF 417
```

Search completed: September 12, 2001, 07:42:20
Job time: 214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2001, 07:33:15 : Search time 20.48 Seconds
(without alignments)
1237.345 Million cell updates/sec

Title: 09-600714-2
Perfect score: 2125
Sequence: 1 MSCKRPSVGRCLPLCALTL.....AKYFDQYFMKPHLAVGFX 418

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_0601:*

- 1: /SIDSB/gcgdata/geneseq/AA1980.DAT:*
- 2: /SIDSB/gcgdata/geneseq/AA1981.DAT:*
- 3: /SIDSB/gcgdata/geneseq/AA1982.DAT:*
- 4: /SIDSB/gcgdata/geneseq/AA1983.DAT:*
- 5: /SIDSB/gcgdata/geneseq/AA1984.DAT:*
- 6: /SIDSB/gcgdata/geneseq/AA1985.DAT:*
- 7: /SIDSB/gcgdata/geneseq/AA1986.DAT:*
- 8: /SIDSB/gcgdata/geneseq/AA1987.DAT:*
- 9: /SIDSB/gcgdata/geneseq/AA1988.DAT:*
- 10: /SIDSB/gcgdata/geneseq/AA1989.DAT:*
- 11: /SIDSB/gcgdata/geneseq/AA1990.DAT:*
- 12: /SIDSB/gcgdata/geneseq/AA1991.DAT:*
- 13: /SIDSB/gcgdata/geneseq/AA1992.DAT:*
- 14: /SIDSB/gcgdata/geneseq/AA1993.DAT:*
- 15: /SIDSB/gcgdata/geneseq/AA1994.DAT:*
- 16: /SIDSB/gcgdata/geneseq/AA1995.DAT:*
- 17: /SIDSB/gcgdata/geneseq/AA1996.DAT:*
- 18: /SIDSB/gcgdata/geneseq/AA1997.DAT:*
- 19: /SIDSB/gcgdata/geneseq/AA1998.DAT:*
- 20: /SIDSB/gcgdata/geneseq/AA1999.DAT:*
- 21: /SIDSB/gcgdata/geneseq/AA2000.DAT:*
- 22: /SIDSB/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1982	93.3	417	21	AA03341
2	1980	93.2	417	20	AA24056
3	1859	87.5	417	21	AA03339
4	1854	87.2	417	21	AA03338
5	1838	86.5	417	21	AA03342
6	1833	86.3	417	21	AA03340
7	497	23.4	473	21	AA042135
8	497	23.4	479	21	AA029656
9	142	6.7	308	21	AA086502
10	109.5	5.2	495	21	AA039787
11	109.5	5.2	577	21	AA039786

12	104.5	4.9	449	19	AA098770	H. pylori GHPO 113
13	104.5	4.9	497	18	AA020821	H. pylori cell env
14	104	4.9	449	21	AA081556	Streptococcus pneu
15	104	4.9	465	22	AA078999	C. glutamicum SRT
16	102.5	4.8	449	19	AA011065	H. pylori ORF 09cp
17	102.5	4.8	1191	21	AA046051	Arabidopsis thalia
18	102.5	4.8	1195	21	AA046050	Arabidopsis thalia
19	102.5	4.8	1325	21	AA046049	Arabidopsis thalia
20	101.5	4.8	461	20	AA038791	Neisseria meningit
21	101	4.8	181	21	AA086279	Human secreted pro
22	100.5	4.7	461	20	AA038792	Neisseria meningit
23	100.5	4.7	758	22	AA079584	Cornebacterium gl
24	100	4.7	348	20	AA019899	B. burgdorferi ant
25	100	4.7	383	20	AA019898	B. burgdorferi ant
26	100	4.7	437	21	AA093270	Amino acid sequenc
27	100	4.7	565	22	AA076817	Cornebacterium gl
28	99	4.7	439	21	AA049911	Arabidopsis thalia
29	99	4.7	442	21	AA013571	Arabidopsis thalia
30	99	4.7	442	21	AA017244	Arabidopsis thalia
31	99	4.7	442	21	AA045471	Arabidopsis thalia
32	99	4.7	442	21	AA045499	Arabidopsis thalia
33	99	4.7	521	21	AA049910	Arabidopsis thalia
34	99	4.7	672	21	AA036170	Novel human transp
35	98.5	4.6	388	19	AA098618	H. pylori GHPO 231
36	98	4.6	405	22	AA079001	C. glutamicum SRT
37	98	4.6	437	21	AA093294	Amino acid sequenc
38	98	4.6	462	20	AA097820	Staphylococcus aur
39	97.5	4.6	437	21	AA049886	Arabidopsis thalia
40	97.5	4.6	476	21	AA070157	Staphylococcus aur
41	97.5	4.6	504	21	AA020282	Arabidopsis thalia
42	97.5	4.6	504	21	AA049885	Arabidopsis thalia
43	97.5	4.6	516	21	AA020281	Arabidopsis thalia
44	97.5	4.6	516	21	AA049884	Arabidopsis thalia
45	97.5	4.6	528	21	AA020280	Arabidopsis thalia

ALIGNMENTS

RESULT	1
AA03341	standard; Protein: 417 AA.
ID	AA03341
XX	
AC	AA03341;
XX	
DF	26-SEP-2000 (first entry)
XX	
DE	Human Rhd protein.
XX	
KW	Human; rhesus blood group system; Rh; RHD; RHCE; sickle cell disease;
KW	thalassaemia; Rhc; Rho; RHE; RHO; alloimmunisation prevention;
KW	autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
KW	vaccine.
XX	
OS	Homo sapiens.
XX	
PN	W0200032632-A2.
XX	
PD	08-JUN-2000.
XX	
PF	01-DEC-1999; 99WO-GB04027.
XX	
PR	01-DEC-1998; 98GB-0026378.
XX	
PA	(UYAB-) UNIV ABERDEEN.
XX	(COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
PI	Urbanlak SJ, Barker RN;
XX	
DR	WPI; 2000-412291/35.
XX	
PT	Composition for prevention of alloimmunization or immunosuppression of
XX	a response elicited by alloimmunization or an autoimmune hemolytic

PT disease, comprises an epitope of a rhesus protein

XX

PS Disclosure: Page 88-90; 92pp; English.

CC Human blood contains the rhesus (Rh) blood group system, and humans can

CC either be RhD positive or negative. This can lead to complications

CC during transfusions or pregnancy if RhD negative individuals are exposed

CC to RhD positive blood, leading to them becoming immunised to produce

CC anti-D. The present invention relates to new human allo- and

CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from

CC RHD, RHC, RHE and RHF proteins. These epitopes bind to T-cells to

CC elicit an immune response, i.e. immunisation. These epitopes can be used

CC as a vaccine for the prevention of alloimmunisation or immunosuppression

CC of a response elicited by alloimmunisation or an autoimmune haemolytic

CC disease. Examples of autoimmune haemolytic diseases are sickle cell

CC disease and thalassemia. The present sequence is the human RhD protein

CC from which the epitopes in the present invention were derived.

XX

SQ Sequence 417 AA;

Query Match 93.3%; Score 1982; DB 21; Length 417;

Best Local Similarity 95.2%; Pred. No. 6.6e-185;

Matches 397; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDOKGLVASVQVGDLTVMMAI 60

1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqvgdlvtmaai 60

DB 61 GLGFLTSSFRHRHSSVAENLFMLALGVOMAILLDGFLSQPSGKVVITLFSTLWTMSA 120

61 glgfltsfrhrhssvafnlfmlalgvomailldgflsqpsgkvvitlfsrlatmsa 120

DB 121 LSVLISVDVAVGLKVNLAOLVYVWLVEVTDGKRLMVISNFFNTHNNMHIVFAAYFG 180

121 lsvlisdvavglkvnlaolvvwlvevtdgkrlmvisnffnthnnmhivfaayfg 180

DB 121 LSVLISVDVAVGLKVNLAOLVYVWLVEVTDGKRLMVISNFFNTHNNMHIVFAAYFG 180

121 lsvlisdvavglkvnlaolvvwlvevtdgkrlmvisnffnthnnmhivfaayfg 180

QY 181 LTVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFPSPVNSALLRSPRIERKNAVFN 240

181 ltvamclpkplpgtedndqatipslsamlgalfwmfpvsfnsallrspierknafvn 240

DB 181 LSVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFPSPVNSALLRSPRIERKNAVFN 240

181 lsvamclpkplpgtedndqatipslsamlgalfwmfpvsfnsallrspierknafvn 240

QY 241 TYRAVAVSVVTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWLPVIGLV 300

241 tyraavsvvtaisgsslahpqckisktyghsavnldpgvavdtschlipspwlamvlglv 300

DB 241 TYRAVAVSVVTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWLPVIGLV 300

241 tyraavsvvtaisgsslahpqckisktyghsavnldpgvavdtschlipspwlamvlglv 300

QY 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYIVLVLDTVGAGNGMIG 360

301 aglisygakylpgccnrvigiphsstimgynfslglleetiylvldtvagngmig 360

DB 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYIVLVLDTVGAGNGMIG 360

301 aglisygakylpgccnrvigiphsstimgynfslglleetiylvldtvagngmig 360

QY 361 FOVLSTIGELSLAIVIALTSGLTALLLNKIRKAPHEAKYFDDQVFWKFPHLAVGF 417

361 fgvllstigelslaivialtsglltllnklrkapheakyfddqvfwkfpahlavgf 417

DB 361 FOVLSTIGELSLAIVIALTSGLTALLLNKIRKAPHEAKYFDDQVFWKFPHLAVGF 417

361 fgvllstigelslaivialtsglltllnklrkapheakyfddqvfwkfpahlavgf 417

RESULT 2

AA924056 standard; Protein: 417 AA.

ID AA924056;

AC AA924056;

DT 04-OCT-1999 (first entry)

XX Protein encoded by the prevalent allele of the RhD gene.

DE

XX Allele: Rhesus D antigen; RhD; weak D phenotype; blood transfusion.

XX Homo sapiens.

OS

XX

PN WO9337763-A2.

XX

PD 29-JUL-1999.

XX

PF 18-DEC-1998; 98WO-EF08319.

XX

PR 23-JAN-1998; 98EP-0101203.

XX

PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTTEMBERG.

XX

PI Flegel WA, Wagner FF;

XX

DR WPI: 1999-469127/39.

DR N-PSDB: AAX86522.

XX

PT Nucleic acid sequences correlated with Rhesus weak D phenotype,

PT useful for screening blood from donors and recipients for

PT transfusion methods

XX

PS Disclosure: Fig 2; 64pp; English.

XX

CC The present sequence is encoded by the prevalent allele of the Rhesus D

CC (RhD) antigen gene. The specification describes a RhD contributing to

CC or indicative of the weak D phenotype, where the RhD polynucleotide

CC carries at least one missense mutation as compared to the wild-type RhD,

CC in its transmembrane and/or intracellular regions, especially in amino

CC acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso

CC that the D antigen does not carry a single missense mutation leading to

CC a F223V or T283I substitution. The probes and antibodies are useful in

CC the methods for detection of weak D phenotypes. Red blood cells, from

CC probands, are useful for the assessment of the affinity, avidity and/or

CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or

CC of anti-globulin or anti-human-globulin antisera. Detecting the presence

CC of the RhD associated with weak D phenotype is useful for determining

CC that a patient in need of a blood transfusion is to be transfused with

CC RhD negative blood from a donor. Alternatively, testing for weak D

CC phenotype RhD in the blood of a donor is useful for determining whether

CC the donor blood should be excluded for transfusion to patients having

CC wild type RhD or weak D types, other than that of the donor weak D

CC type.

XX

SQ Sequence 417 AA;

Query Match 93.2%; Score 1980; DB 20; Length 417;

Best Local Similarity 95.2%; Pred. No. 1e-184;

Matches 397; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 1 MSCKYPSVGRCLPLCALTEALILLFYFTHYDASLEDOKGLVASVQVGDLTVMMAI 60

1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqvgdlvtmaai 60

DB 1 msckypsvgrclplcaltealilfffthydasldeqkglvasyqvgdlvtmaai 60

DB 61 GLGFLTSSFRHRHSSVAENLFMLALGVOMAILLDGFLSQPSGKVVITLFSTLWTMSA 120

61 glgfltsfrhrhssvafnlfmlalgvomailldgflsqpsgkvvitlfsrlatmsa 120

DB 61 GLGFLTSSFRHRHSSVAENLFMLALGVOMAILLDGFLSQPSGKVVITLFSTLWTMSA 120

61 glgfltsfrhrhssvafnlfmlalgvomailldgflsqpsgkvvitlfsrlatmsa 120

QY 121 LSVLISVDVAVGLKVNLAOLVYVWLVEVTDGKRLMVISNFFNTHNNMHIVFAAYFG 180

121 lsvlisdvavglkvnlaolvvwlvevtdgkrlmvisnffnthnnmhivfaayfg 180

DB 121 LSVLISVDVAVGLKVNLAOLVYVWLVEVTDGKRLMVISNFFNTHNNMHIVFAAYFG 180

121 lsvlisdvavglkvnlaolvvwlvevtdgkrlmvisnffnthnnmhivfaayfg 180

QY 181 LTVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFPSPVNSALLRSPRIERKNAVFN 240

181 ltvamclpkplpgtedndqatipslsamlgalfwmfpvsfnsallrspierknafvn 240

DB 181 LSVAMCLPKPLPGTEEDNDQATIPSLSAMLGALFLMFPSPVNSALLRSPRIERKNAVFN 240

181 lsvamclpkplpgtedndqatipslsamlgalfwmfpvsfnsallrspierknafvn 240

QY 241 TYRAVAVSVVTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWLPVIGLV 300

241 tyraavsvvtaisgsslahpqckisktyghsavnldpgvavdtschlipspwlamvlglv 300

DB 241 TYRAVAVSVVTAISGSSLAHPQCKISKTYGHSANVDPGVAVDTSCHLIPSPWLPVIGLV 300

241 tyraavsvvtaisgsslahpqckisktyghsavnldpgvavdtschlipspwlamvlglv 300

QY 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYIVLVLDTVGAGNGMIG 360

301 aglisygakylpgccnrvigiphsstimgynfslglleetiylvldtvagngmig 360

DB 301 AGLISVYGAKYLPGCCNRVIGIPHSSTMGYNFSLGLLEETIYIVLVLDTVGAGNGMIG 360

301 aglisygakylpgccnrvigiphsstimgynfslglleetiylvldtvagngmig 360

QY 361 FOVLSTIGELSLAIVIALTSGLTALLLNKIRKAPHEAKYFDDQVFWKFPHLAVGF 417

361 fgvllstigelslaivialtsglltllnklrkapheakyfddqvfwkfpahlavgf 417

DB 361 FOVLSTIGELSLAIVIALTSGLTALLLNKIRKAPHEAKYFDDQVFWKFPHLAVGF 417

361 fgvllstigelslaivialtsglltllnklrkapheakyfddqvfwkfpahlavgf 417

RESULT 3

1D	AAB03339	standard; Protein; 417 AA.
XX	AC	AAB03339;
XX	DT	26-SEP-2000 (first entry)
XX	DE	Human Rhce protein.
XX	KW	Human; rhesus blood group system; Rh: Rhd; RHCE; sickle cell disease; thalassemia; RHC: Rhc; RHe; Rhd; alloimmunisation prevention; autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive; vaccine.
XX	OS	Homo sapiens.
XX	PN	WO200032632-A2.
XX	PD	08-JUN-2000.
XX	PF	01-DEC-1999; 99MO-GB04027.
XX	PR	01-DEC-1998; 98GB-0026378.
XX	PA	(UYAB-) UNIV ABERDEEN.
XX	PA	(COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX	PI	Urbanlak SJ, Barker RN;
XX	DR	WPI: 2000-412291/35.
XX	PT	Composition for prevention of alloimmunization or immunosuppression of a response elicited by alloimmunization or an autoimmune hemolytic disease, comprises an epitope of a rhesus protein -
XX	PS	Disclosure; Page 85-86; 92pp; English.
XX	CC	Human blood contains the rhesus (Rh) blood group system, and humans can either be RhD positive or negative. This can lead to complications during transfusions or pregnancy if RhD negative individuals are exposed to RhD positive blood, leading to them becoming immunised to produce anti-D. The present invention relates to new human allo- and auto-reactive T-cell epitopes (AA99760-99769 and AAB03201-B03337) from RhD, RHC, Rhd, RHe and Rhe proteins. These epitopes bind to T-cells to elicit an immune response, i.e. immunisation. These epitopes can be used as a vaccine for the prevention of alloimmunisation or immunosuppression of a response elicited by alloimmunisation or an autoimmune haemolytic disease. Examples of autoimmune haemolytic diseases are sickle cell disease and thalassemia. The present sequence is the human Rhce protein from which the epitopes in the present invention were derived.
SO	Sequence	417 AA;
QY	Query Match	87.5%; Score 1859; DB 21; Length 417;
	Best Local Similarity	89.9%; Pred. No. 6, 4e-173;
	Matches	375; Conservative 10; Mismatches 32; Indels 0; Gaps
QY	1	MSCKPRSVGRCPLCALFTLEALILLYFFPTHYDASLEDOKGLVASVOYOGDLYVMAAI 60
DB	1	mskkyprsvrrrcplcalctleaaallllyffthdaaleqkglvasygyqdlvmaai 60
QY	61	GIGFLTTSFRHSMSSVAFNFMALGVQMAIILLDGFLSPFSGKVIYTLFTSLATMSEA 120
DB	61	gigfltsfrhssvvaflnfmalgvqaailldgflsqfsgkxviltfslrlatmsea 120
QY	121	LSVILSVADVAVLGNKIALQOLVYMLVEYTDGNTLRMTVSNIFNDYDHHNMHIVFAAYFG 180
DB	121	msvilsaadvavlgkynlqglvymvlveretalgltlmvsnlfnctdyhnmnlrhfyvfaayfg 180
QY	181	LTAVAMCKPKPLPEGETEDNDQATATPISLSAMIGALFTLMEF-PVSNALSRLSPIERKNAVEN 240
DB	181	ltvavcmkpkplpketddqgratpissamigaltfmlmfvpsvnsallfspdigrknamfn 240

```

Oy 241 TTYAAVAVSWIRIASSSLAHFQGRISKRTYGHSAVLPEGVAVDTSHLIPSWLPIVLGLV 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TTYAAVAVSWIRIASSSLAHFQGRISKRTYGHSAVLPEGVAVDTSHLIPSWLPIVLGLV 300
    TTYAAVAVSWIRIASSSLAHFQGRISKRTYGHSAVLPEGVAVDTSHLIPSWLPIVLGLV 300
Oy 301 AGLISVRCAKYLPGCCNENVLGIPHSISGVNPSLLGLEETIITYVLAVLDIVGAGNGMIG 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 AGLISVRCAKYLPGCCNENVLGIPHSISGVNPSLLGLEETIITYVLAVLDIVGAGNGMIG 360
    AGLISVRCAKYLPGCCNENVLGIPHSISGVNPSLLGLEETIITYVLAVLDIVGAGNGMIG 360
Oy 361 FOVLISIGLSLAVIALTSGLTALNLNLRKAPHEAKYFDQVFWKPHLAVGF 417
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 FOVLISIGLSLAVIALTSGLTALNLNLRKAPHEAKYFDQVFWKPHLAVGF 417
    FGVLLISIGLSLAVIALTSGLTALNLNLRKAPHEAKYFDQVFWKPHLAVGF 417

RESULT 4
ID AAB03338 standard; Protein: 417 AA.
XX AAB03338;
XX AAB03338;
XX DT 26-SEP-2000 (first entry)
XX DE Human RhCE protein.
XX KW Human; rhesus blood group system; Rh, RHD, RhCE; sickle cell disease;
XX KW Thalassaemia; RHC; Rhc; RHE; Rhe; alloimmunisation prevention;
XX KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
XX KW vaccine.
XX OS Homo sapiens.
XX PN WO200032632-A2.
XX PD 08-JUN-2000.
XX PE 01-DEC-1999; 99WO-GB04027.
XX PR 01-DEC-1998; 98GB-0026378.
XX PA (UYAB-) UNITV ABERDEEN.
XX PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX PI Urbanlak SJ, Baker RM.
XX DR WPI: 2000-412291/35.
XX PT Composition for prevention of alloimmunization or immunosuppression of
XX PR a response elicited by alloimmunization or an autoimmune hemolytic
XX PR disease, comprises an epitope of a rhesus protein
XX PS Disclosure: Page 83-84; 92pp; English.
XX CC Human blood contains the rhesus (Rh) blood group system, and humans can
XX CC either be RhD positive or negative. This can lead to complications
XX CC during transfusions or pregnancy if RhD negative individuals are exposed
XX CC to RhD positive blood, leading to them becoming immunised to produce
XX CC anti-D. The present invention relates to new human allo- and
XX CC auto-reactive T-cell epitopes (AAV99760-Y99769 and AAB03201-B03337) from
XX CC Rho, RHC, Rhc, RHE and Rhe proteins. These epitopes bind to T-cells to
XX CC elicit an immune response, i.e. immunisation. These epitopes can be used
XX CC as a vaccine for the prevention of alloimmunisation or immunosuppression
XX CC of a response elicited by alloimmunisation or an autoimmune haemolytic
XX CC disease. Examples of autoimmune haemolytic diseases are sickle cell
XX CC disease and Thalassaemia. The present sequence is the human RhCE protein
XX CC from which the epitopes in the present invention were derived.
XX SQ Sequence 417 AA;

Query Match 87.2%; Score 1854; DB 21; Length 417;
Best Local Similarity 89.7%; Pred. No. 2e-172;
Matches 374; Conservative 10; Mismatches 33; Indels 0; Gaps 0;

```

OY 1 MSKPRSVGVCPLPCALTEALAILLEFPEETHDALEDOCKGVASXYOGDLTWAAI 60
 Db 1 MSKPRSVVRCPLPCALTEALAILLEFPEETHDALEDOCKGVASXYOGDLTWAAI 60
 OY 61 GIGELTSSFRHSSWSVAFNLFMLALGVOMAILLDCFLSQFSPSGKVYTFLESTWLATMSA 120
 Db 61 GIGELTSSFRHSSWSVAFNLFMLALGVOMAILLDCFLSQFSPSGKVYTFLESTWLATMSA 120
 OY 121 LSVLISDVALGKYNLAQLVVMVLEVTDLGNLRMTSNIFNTEDYHNMNMHIYVFAAYFG 180
 Db 121 LSVLISDVALGKYNLAQLVVMVLEVTDLGNLRMTSNIFNTEDYHNMNMHIYVFAAYFG 180
 OY 181 LTVAMCPLKPLPEGEEDDDQFATIPRSISAMGALFIMMFPSVNSALLRSPRIERKNAVEN 240
 Db 181 LTVAMCPLKPLPEGEEDDDQFATIPRSISAMGALFIMMFPSVNSALLRSPRIERKNAVEN 240
 OY 241 TYVAAVSVVMAIAGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
 Db 241 TYVAAVSVVMAIAGSSSLAHPOGKISKTYGHSAYLPEGVAVDTSCHLIPSPMLPIVLGLV 300
 OY 301 AGLSIVTGAKYLPCCCNRVLGIPHSSIMGYNFSLLGLLEETIYVILVLDTVGAGNGMGIG 360
 Db 301 AGLSIVTGAKYLPCCCNRVLGIPHSSIMGYNFSLLGLLEETIYVILVLDTVGAGNGMGIG 360
 OY 361 FOYLSTGELSLAVIYVIAITSGSLTALLNLKIRAPRPAKAYFPDQVPMKPRHLAVGF 417
 Db 361 FOYLSTGELSLAVIYVIAITSGSLTALLNLKIRAPRPAKAYFPDQVPMKPRHLAVGF 417
 RESULT 5
 ID AAB03342 standard; Protein: 417 AA.
 AC AAB03342;
 DT 26-SEP-2000 (first entry)
 DE Human Rhce protein.
 XX Human, rhesus blood group system; Rh; RhD; RhCE; sickle cell disease;
 KW thalasassaemia; RHC; Rhc; RHE; Rho; alloimmunisation prevention;
 KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
 KW vaccine.
 OS Homo sapiens.
 PN MO200032632-A2.
 PD 08-JUN-2000.
 XX 01-DEC-1999; 99MO-GB04027.
 PR 01-DEC-1998; 98GB-0026378.
 XX (UYAB-) UNIV ABERDEEN.
 PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
 PI Urbanlak SJ, Barker RN;
 DR WP1: 2000-412291/35.
 XX
 CC Human blood contains the rhesus (Rh) blood group system, and humans can
 CC either be RhD positive or negative. This can lead to complications
 CC during transfusions or pregnancy if RhD negative individuals are exposed
 CC to RhD positive blood, leading to them becoming immunised to produce
 CC anti-D. The present invention relates to new human allo- and
 CC auto-reactive T-cell epitopes (AAY9760-Y9769 and AAB03201-B03347) from

[illegible]

XX WPI; 2000-412291/35.
XX Composition for prevention of alloimmunization or an autoimmune hemolytic
PT response elicited by alloimmunization or an autoimmune hemolytic
PT disease, comprises an epitope of a Rhesus protein
XX
PS Disclosure; Page 86-88; 92pp; English.
XX Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AA99760-Y99769 and AAB03201-B03337) from
CC Rhd, Rhc, Rho, and Rhr proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune hemolytic
CC disease. Examples of autoimmune hemolytic diseases are sickle cell
CC disease and thalassemia. The present sequence is the human Rhesus protein
CC from which the epitopes in the present invention were derived.
XX
SQ Sequence 417 AA;

Query Match 86.3%; Score 1833; DB 21; Length 417;
Best Local Similarity 88.7%; Pred. No. 2.2e-170;
Matches 370; Conservative 12; Mismatches 35; Indels 0; Gaps 0;

QY 1 MSCKRPSVGRCLPLCALTEALILFEFTHYDASLEQKGLVASVOGDLTVMAAI 60
DB 1 MSKYPSTVRCPLWALTLEAIIILFFTHYDASLEQKGLVASVQGDLTMAAI 60

QY 61 GLGELTSSFRHSSVAFNFMALGVQVAILDGLSPSPSKVYTLFSTLATHSA 120
DB 61 GLGELTSSFRHSSVAFNFMALGVQVAILDGLSPSPSKVYTLFSTLATHSA 120

QY 121 LSVLISDAVAGKYNLAOLVYVLEVTDLGNLBMVSNINFTYHMMHIIYFAAYFG 180
DB 121 LSVLISDAVAGKYNLAOLVYVLEVTDLGNLBMVSNINFTYHMMHIIYFAAYFG 180

QY 121 MSVLSAGAVLGKYNLAOLVYVLEVTDLGNLBMVSNINFTYHMMHIIYFAAYFG 180
DB 121 MSVLSAGAVLGKYNLAOLVYVLEVTDLGNLBMVSNINFTYHMMHIIYFAAYFG 180

QY 181 LTVAMCLPKPLPTEGTDNDQATIPSLSAMGALFLMWFEPVSVALLSPIERKNAYFN 240
DB 181 LTVAMCLPKPLPTEGTDNDQATIPSLSAMGALFLMWFEPVSVALLSPIERKNAYFN 240

QY 181 LTVAMCLPKPLPTEGTDNDQATIPSLSAMGALFLMWFEPVSVALLSPIERKNAYFN 240
DB 181 LTVAMCLPKPLPTEGTDNDQATIPSLSAMGALFLMWFEPVSVALLSPIERKNAYFN 240

QY 241 TYRVAVSVYTAIGSSLAHQKISTKYGSAVLPQGVAVTSCHLIPSPWLPVIGLV 300
DB 241 TYRVAVSVYTAIGSSLAHQKISTKYGSAVLPQGVAVTSCHLIPSPWLPVIGLV 300

QY 241 TYRVAVSVYTAIGSSLAHQKISTKYGSAVLPQGVAVTSCHLIPSPWLPVIGLV 300
DB 241 TYRVAVSVYTAIGSSLAHQKISTKYGSAVLPQGVAVTSCHLIPSPWLPVIGLV 300

QY 301 AGLISVAGAKYLPCCNRVIGIPHSIMGVNFSLLGLLEITIVLVLDVVGANGMIG 360
DB 301 AGLISVAGAKYLPCCNRVIGIPHSIMGVNFSLLGLLEITIVLVLDVVGANGMIG 360

QY 301 AGLISVAGAKYLPCCNRVIGIPHSIMGVNFSLLGLLEITIVLVLDVVGANGMIG 360
DB 301 AGLISVAGAKYLPCCNRVIGIPHSIMGVNFSLLGLLEITIVLVLDVVGANGMIG 360

QY 361 FOVLISIGELSLAIVALTSGTLTALDLNLRKAPHEAKYFDQVWEPKPELAVGF 417
DB 361 FOVLISIGELSLAIVALTSGTLTALDLNLRKAPHEAKYFDQVWEPKPELAVGF 417

RESULT 7
AAB42135
ID AAB42135 standard; Protein: 473 AA.
XX
AC AAB42135;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1899 polypeptide sequence SEQ ID NO:3798.
XX
KW Human; open reading frame; ORFX: detection; cytostatic; hepatotropic;
KW vulnary; antiparkinsonian; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteoprotic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-0508621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shlmkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX
XX N-PSDB; AAC76344.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 2948-2949; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX CC antiparkinsonian; antiparkinsonian; neurotropic; neuroprotective;
XX CC osteoprotic; anticonvulsant; antirheumatic; immunosuppressant;
XX CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX CC antidiabetic; hypotensive; dermatological; immunosuppressive;
XX CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX CC antihypertoid; and antianemic. The sequences can be used for determining
XX CC the presence of or predisposition to, or preventing or treating
XX CC pathological conditions associated with an ORFX-associated disorder. The
XX CC nucleic acids can be used to express ORFX proteins in gene therapy
XX CC vectors. The proteins and nucleic acids may be used to treat cancers,
XX CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX CC graft vs host disease, cardiovascular disease, diabetes mellitus,
XX CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX
SQ Sequence 473 AA;

Query Match 23.4%; Score 497; DB 21; Length 473;
Best Local Similarity 29.4%; Pred. No. 4.4e-40;
Matches 129; Conservative 77; Mismatches 183; Indels 50; Gaps 8;

QY 13 LPLCALTEALILFEFTHYD-----ASLEQKGLVASVOGDLTVMA 58
DB 5 LPLCALTEALILFEFTHYD-----ASLEQKGLVASVOGDLTVMA 58

QY 59 AIGGELTSSFRHSSVAFNFMALGVQVAILDGLSPSPSKVYTLFSTLATHSA 118
DB 59 AIGGELTSSFRHSSVAFNFMALGVQVAILDGLSPSPSKVYTLFSTLATHSA 118

QY 65 FVGLISIGELSLAIVALTSGTLTALDLNLRKAPHEAKYFDQVWEPKPELAVGF 124
DB 65 FVGLISIGELSLAIVALTSGTLTALDLNLRKAPHEAKYFDQVWEPKPELAVGF 124

PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139753.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142380.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144351.
PR 19-JUL-1999; 9905-0144352.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.

PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148365.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151338.
PR 01-SEP-1999; 9905-0151330.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161820.
PR 28-OCT-1999; 9905-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 5.2%; Score 109.5; DB 21; Length 495;
Best Local Similarity 19.8%; Pred. No. 0.028;
Matches 97; Conservative 77; Mismatches 160; Indels 157; Gaps 23;
QY 12 CLPLCALTEALILFFFTHTDASLEDDKGLVASYQVODLTVMALIGLFTSSRR 71
| | | | | : : : : : | | | | | : : : : :
DB 7 cdplcsvdeps---syfeanygpktdilkala|---laaalgtlaa|-----n 49
QY 72 HSNSSVAFNLFM-LALGVOMA-IILDDGLSOPSGRVITFSTWLA-TMSALSVLSV- 127
| | | | | : : : : : | | | | | : : : : :
DB 50 hswvaangdvamallfigiyagiiifeeslafnsgidilmaavslwvrsigapstela1 109
QY 128 ----DAVLGRVNLQQLVWVLEVEVTL-GNLRWVSNIFNTDYHMMHMYFAAYF-- 179
| | | | | : : : : : | | | | | : : : : :
DB 110 dlgataevseivfflllgamtiveidabgfkivtdnl-cttrpkclllwvgfvtflls 168
QY 180 -----GLTVAMCLRPRLPEGETEDDQATIPSLSAMGALF-----LW---- 217
| | | | | : : : : : | | | | | : : : : :
DB 169 sildnlstlwmvslklklypqseyrnllygvvvaanaaggawtpigdvltlmlwhgq 228
QY 218 -----MERPSVNGALLRSPIERKNAYENTYVAVAVSVTAISG-----SSL 258
| | | | | : : : : : | | | | | : : : : :
DB 229 tscptnkdflfsvvs--lavpl-----almstsevnkgkqdkvlassek 274
QY 259 AHPOGKISKITYGHSVAVLPEGVAAVTSCHLIPSPWLPVLGVAGLISVYGAKYLPGCCNR 318
| | | | | : : : : : | | | | | : : : : :
DB 275 maprgklivfgvgalav--fypvfkaltglp-pymgilllg----- 312
QY 319 VLGIPIHSIMGYNSLLGLEETIYF-----VLIVDTVGAGNGMIGFOVLLS 366
| | | | | : : : : : | | | | | : : : : :
DB 313 -----gvwlwlttdalhygeserqklkvpqalsrldtga--lfflgllls 355
QY 367 IGETSLAIY-----IATLSGLTALLNLNKIRKAP---HEAKYF-DDQ 405
| | | | | : : : : : | | | | | : : : : :
DB 356 vssleaaqgilrelanyldanipveliasaigvvasaiidhvpivaatcmgmdltsfpqds 415
QY 406 VFMKFPHLAVG 416
| | | | | : : : : :
DB 416 efwqliafcag 426
RESULT 11
AAG39786
ID AAG39786 standard; Protein; 577 AA.
XX
AC AAG39786;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49280.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143942.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

AAW98770;
 31-MAR-1999 (first entry)
 H. pylori GHPO 1130 protein.
 GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 peptic ulcer disease.
 Helicobacter pylori.
 WO9843478-A1.
 08-OCT-1998.
 01-APR-1998; 98WO-US06371.
 29-JUL-1997; 97US-0902615.
 01-APR-1997; 97US-0833457.
 24-JUN-1997; 97US-0881227.
 (HUMA-) HUMAN GENOME SCI INC.
 (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 Al-Garawi A, Kleantous H, Miller C, Comen RP, Tomb J;
 WPI; 1998-542293/46.
 N-PSDB; AAX14489.
 New isolated Helicobacter polynucleotides - used to develop products
 for the diagnosis, prevention and treatment of Helicobacter
 infections and gastrointestinal diseases
 Claim 8; Page 1663-1665; 2054pp; English.
 This sequence represents a Helicobacter pylori GHPO protein of the
 invention. The polypeptides can be used for preventing or treating
 Helicobacter infections, and gastroduodenal diseases associated with
 these infections, including acute, chronic, and atrophic gastritis, and
 peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 used for the production of antibodies. The products can also be used for
 detection and diagnosis.

Sequence 449 AA;

Query Match 4.9%; Score 104.5; DB 19; Length 449;
 Best Local Similarity 20.8%; Pred. No. 0.076;
 Matches 85; Conservative 63; Mismatches 114; Indels 147; Gaps 20;

27 LFFEFTHYDASLEDDOKGLVASYGVODLTVMMAIGLFTSFRHRSWSSVAFNLEMLAL 86
 130 llyff--lhnatpsqngf-----glpmatclafalgvmllygr--vptalkvlltl 178
 87 GYVMAILLDGLFLOPSGKVVITLF-----SIWATMSALSVLISDAVIGKVNLAOLV 140
 179 av-----addlgavivalfytnlkfawl--lgalgvl-vlavlnlrmssl 225
 141 VMVLVEYTDGNNRMWYSNFNFDYHNMNMHTYVFAAYFGLVAMCLPKLPDESTED--- 197
 226 pylllyv-----llwfcvbgsglhatfaa---vllaflmipvkxpksknlvel 269
 198 -----ndoratisLSAAMLGALFLMFR-PSVNSALRSPIERKNVFN 240
 270 lelgkyraetsgallskqegelhshleekasa-----lqspier----- 309
 241 TYVAVAVSVTATISGSLAHPOGRISKTYGHSALP-----EGVAVDTSCHLIPSEWLP 294
 310 -----lehlfiaplsygf-----lmpifafanagvsdsslnlevd---k 345
 295 IYGLVAGLISYGAAYLPGCCNRVLGPHSSIMGYNFSLGLLEETIYIVLVLDIVGA 354
 346 vllgvllgl-----clgkplgllflitfiseklkltarpkgiswvhl-----ga 389

QY 355 G-NGMICEVYLLSIGELS-----LAIVI-ALRTSGLLTALL 388
 390 gllagigftmsmflsnlafscenhkdamevakiallglslisgligaly 438

RESULT 13
 AAW20821
 ID AAW20821 standard; Protein; 497 AA.
 XX
 AC AAW20821;
 16-JUL-1997 (first entry)
 H. pylori cell envelope inner membrane protein 11cell1603orf16.
 Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 Helicobacter pylori.
 WO9640893-A1.
 19-DEC-1996.
 06-JUN-1996; 96WO-US09122.
 01-APR-1996; 96US-0630405.
 07-JUN-1995; 95US-0487032.
 (ASTR) ASTRA AB.
 Berglundh OF, Smith D, Mellgaerd BL;
 WPI; 1997-052306/05.
 N-PSDB; AAT68074.
 Helicobacter pylori nucleic acid sequences and related
 polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 infection, and to detect Helicobacter
 Claim 56; Page 1226-1227; 1481pp; English.

The present sequence shows a Helicobacter pylori cell envelope
 inner membrane protein.
 The protein may be used in a vaccine to prevent or treat H. pylori
 infection or to identify H. pylori polypeptide binding compounds,
 useful as potential H. pylori life cycle activators or inhibitors.
 The genomic sequence of H. pylori (ATCC 55679) was determined from
 overlapping contigs generated by mechanically shearing the bacterial
 DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 and the predicted coding regions defined by computer evaluation. To
 identify likely H. pylori antigens for vaccine development, the amino
 acid sequences predicted from various ORF were analysed for significant
 homology to other known or exported membrane proteins. Having identified
 and determined the sequences of interest, particular regions can be
 isolated from H. pylori by PCR amplification for recombinant polypeptide
 production, e.g. in E. coli hosts.

Sequence 497 AA;

Query Match 4.9%; Score 104.5; DB 18; Length 497;
 Best Local Similarity 22.7%; Pred. No. 0.086;
 Matches 111; Conservative 52; Mismatches 174; Indels 153; Gaps 27;

15 LCAUTTEAAILLFYFFTH---YDASLEDDOKGLVASYGVODL-----TVMAI----- 60
 48 lcvlvsfigavvllfigwhngsyekylfdwi-vvgnfkvgfslmldnlnvmlvvlvs 106
 61 -----GLGFL--TSSRRHRSWSSVAFNLEMLALGVOMAILLDGLFLOPS-----GKVV 108

Db 107 flvhyvsiygmehdgtfnryfyslsgtfvsmlyl-----vlsdnflglfigwvgyjcsy 161
 QY 109 TLFSTVATMSALSVLISVDVAGKVNLAOLVWVLEVTDLGNL----- 153
 Db 162 lllgfwyhhksa-----nnsiaefvtridldlmgllllfwnfgclgy 207
 QY 154 RMVTSNFTNDYHNMNMHIYF-----AAVFGT--LYAWCLPKPLPGEDEDNDQRAF 203
 Db 208 keyfsmInnadysm-lfyisvflfigamgksqgfmhtwlanamegctp----- 255
 QY 204 IPSLSAALGLFLMFPPSVNSALRSPIERKNAVENT--YAAVAVSVVTATSGSLAHAPQ 262
 Db 256 ---vsallth--atmvtaagyvlltrnpl--ysavfevgyfiacigafvalfgasma1vn 308
 QY 263 GKISRTYGHSAVLPEG---VAVDTSC-----HLIPSPMPrlpVLGLVAGL----- 303
 Db 309 kdlkrtvaystlsglgymfraaglgayalafnhfthafksllflfsgnmvhamednd 366
 QY 304 ISVYGAAYLPGCCNRVLGPHSSIM-----GYNFS-----LGLLEETIYVLLVL 349
 Db 369 ltkmgalykpmrltavfmligsva1cgy1pfaay-fskdklleafgmbhhlwfvll-- 425
 QY 350 DYVAGNGMIGFQVLLSIGELSLAIVATLNSGLLTALLNLKIRKAPHEAKY-FDDQYFW 408
 Db 426 -----1g-----alfafysfrllmlvf---fapkdhe1nhpgegkfh 461
 QY 409 KFPPLAV-GF 417
 Db 462 afehatvrgf 471

RESULT 14

AA81556 ID AAY81556 standard; Protein: 449 AA.

XX AAY81556;

XX 24-MAY-2000 (first entry)

XX Streptococcus pneumoniae type 4 protein sequence #56.

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease.

XX Streptococcus pneumoniae.

XX MO200006737-A2.

XX 10-FEB-2000.

XX 27-JUL-1999; 99MO-GH02451.

XX 27-JUL-1998; 98GB-0016337.

XX 19-MAR-1999; 99US-0125164.

XX (MICR-) MICROBIAL TECHNIQS LTD.

XX Gilbert CFG, Hansbro PM;

XX WPI; 2000-195300/17.

PT New Streptococcal protein, useful as a vaccine, for diagnosis of
 PT pneumococcal diseases and for screening agents capable of antagonizing
 PT or inhibiting expression of the protein

XX Claim 1; Page 78; 108pp; English.

CC AAY81501 to AAY81679 represent specifically claimed protein sequences
 CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
 CC The sequences have antibacterial and antiinflammatory properties.
 CC The protein sequences, and fragments of them, are useful as immunogens

CC and/or antigens. The nucleotide sequences can be used in vaccines and in
 CC diagnostic assays. The proteins and nucleotides can be useful for the
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also
 CC useful for screening an agent capable of antagonizing, inhibiting or
 CC interfering with the function or expression of the proteins in which the
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
 CC and meningitis. AA05591 to AA05614 represent primers used in the
 CC exemplification of the present invention.

SO Sequence 449 AA:

Query Match 4.9%; Score 104; DB 21; Length 449;
 Best Local Similarity 22.0%; Pred. No. 0.085;
 Matches 85; Conservative 53; Mismatches 149; Indels 100; Gaps 17;

QY 43 GLVASYOVGDLY--MAALGLGFLTNSFRHSMSSVAFNLFMALGVQAYILLDGLSQ 100
 Db 20 gklagkfivvraltaagnaavftlvgsmflvflspatfsf-----wpivadlfsas 71
 QY 101 FPSGKVYITLFSIWL-----ATMSALSVLISVDVAGKVNLAOLVWVLEVTDLGNLRV 156
 Db 72 fdk-----ftslymwanyatmgs1slyf-----vls1ageltkiyaeel----- 112
 QY 157 ISNIFNTDYHNMNMHIYFPAAY-FGLTVAMCLPKPLPGEDEDNDQAFATPSLSAMLGALF 215
 Db 113 -----nmprlngallalmafvmfv-----pqllfdggmktvtslke--gav1 153
 QY 216 L--WMFPPSVNSALRSPIERKNAVENT--YAAVAVSVVTATSGSLAHAPQKISTY 270
 Db 154 adgwangnvv-----arfgtgcifcalimalavtlvlytmcvkhnw1km--- 197
 QY 271 HSAVLPEGVAVDTSCHLIPSPMPrlpVLGLVAGLISVYGAAYLPGCCNRVLGPHSSIMGY 330
 Db 198 -peavpegvsrgfta-lvpgfvavfvlfnqllvamgtliff-----kvalafgfvsn1 250
 QY 331 NFSLLGLLEETIYVLLVLDTVGAGNGMIGFQVLLSI-----GELSLA 373
 Db 251 tnswi1gm--lly1l1tqllwlg1hganlvfafvsp1alanmaanaaghfavagfsm 308
 QY 374 IVATLNSGLLTALLNLKIRKAPHEAK 400
 Db 309 fva1agsgac1g1c1y1afas1seq1k 335

RESULT 15

AA878999 ID AAB78999 standard; Protein: 465 AA.

XX AAB78999;

XX 30-APR-2001 (first entry)

XX C. glutamicum SRT protein sequence SEQ ID NO:258.

KW Corynebacterium glutamicum; stress; resistance; tolerance; SRT;
 KW fine chemical production; organic acid; pyrimidine base; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW evolutionary study; environmental hazard; fermentation.

XX Corynebacterium glutamicum.

XX WO200100804-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000MO-IB00922.

XX 25-JUN-1999; 99US-0141031.

PR 01-JUL-1999; 99DE-1030429.
 PR 01-JUL-1999; 99US-0142692.

PR 08-JUL-1999; 99DE-1031413.
 PR 08-JUL-1999; 99DE-1031457.
 PR 08-JUL-1999; 99DE-1031541.
 PR 09-JUL-1999; 99DE-1032209.
 PR 09-JUL-1999; 99DE-1032230.
 PR 14-JUL-1999; 99DE-1032914.
 PR 27-AUG-1999; 99DE-1040764.
 PR 27-AUG-1999; 99US-0151214.
 PR 31-AUG-1999; 99DE-1041382.

XX (BAD) BASF AG.
 XX
 XX

PI Pompejus M, Kroege B, Schroeder H, Zelder O, Habener G, Lee H;
 PI Kim H;
 PI

DR MPI: 2001-061972/07.
 DR N-PSDB; AAF71112.

PT New isolated *Corynebacterium glutamicum* nucleic acid encoding a stress,
 PT tolerance or resistance protein, for production or modulation of
 PT production of fine chemicals, such as, e.g. amino acids, lipids,
 PT carbohydrates, or enzymes -
 XX
 XX

PS Claim 20; Page 465-466; 526pp; English.

XX AAF70984 to AAF71133 encode the *Corynebacterium glutamicum* stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The SRT
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (III) encoded by them are used for diagnosing the presence
 CC or activity of *Corynebacterium diptheriae*. (I), (II), (III) and host
 CC cells containing them can be used to map the genomes of organisms related
 CC to C. glutamicum, to identify and localise C. glutamicum sequences of
 CC interest, in evolutionary studies, in determination of SRT protein
 CC regions required for function, in modulating the SRT protein activity,
 CC and in modulating the activity of an SRT pathway. (II) are used to permit
 CC C. glutamicum to survive in an environment that is normally
 CC environmentally or chemically hazardous to it. (I) and protein molecules
 CC encoded by it increase the survival of C. glutamicum to chemical and
 CC environmental hazards and provide a means for continued growth and
 CC multiplication in large scale fermentative growth conditions. By
 CC increasing the growth rate or maintaining a normal growth rate in poor or
 CC toxic conditions, the yield, production and/or efficiency or production
 CC of fine chemicals from a culture may be increased.
 CC
 CC
 XX

SQ Sequence 465 AA:

Query Match 4.9%; Score 104; DB 22; Length 465;
 Best Local Similarity 19.8%; Pred. No. 0.089;
 Matches 79; Conservative 64; Mismatches 134; Indels 122; Gaps 17;

OY 13 LPTCALTL-----EAILILFFFTHYDASLEDOKGLVASYQVQDILTVMATIGLGF 64
 Db 104 vpicllllrgrvnekyall-----gltvsvng--lgyvdalagw 145
 OY 65 L--TSSFRHSMSVAFNLFALGVOMATILDOFISOPSGKVITLFSIWLATMSALS 122
 Db 146 laetlgrfslfwmaaf-cavaalalpfsv-keetaeepk-----mdwl---gvlp 192
 OY 123 VLISVDATL-----GKNVLAQLVNVVLEVTDLGNLRWVISNIFNTDYHNMNHIVFA 176
 Db 193 lavisgslmafeagkllgaanwllvvllflfigagv-----lffynlekrykplvsv 246
 OY 177 AYFGLTAVAKCLPRPLPGSTEDNDQRTATIPSSAMLGALFLMFRFSPVNSALRLRPIERKN 236
 Db 247 eylgrrttwaill-----stlltmtgvfawmgllpnlagdaanga----- 287

OY 237 AVENTYVAVASVVT-----AISGSSLAHPQGIKSTYGHSAVLPEGVAVDTSCHLIP 289
 Db 288 -----gmsasvsvswlltpyalaglvfgplagllagkfgylvlq----- 327
 OY 290 SPWLPVLGLVAGLISVTCAGKYLPGCCNRY-----LGIPHSSTMGVNFSLGLLEET 341
 Db 328 -----lgiactllgvagatflvgstshlaylgslfgytygjanlmnglg----- 375
 OY 342 IYIVLLVLDTVGAGNGMIGFQVLLSIGELSLATVIALTS 380
 Db 376 --lvispanngylypymnagafmlyag-istfallfavst 411

Search completed: September 12, 2001, 07:38:44
 Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 12, 2001, 07:33:51 ; Search time 12.41 Seconds
(without alignments)
693.535 Million cell updates/sec

Title: 09-600714-2
Perfect score: 2125
Sequence: 1 MSCXPRSVGRCLPLCALTL.....AKYPDQVFMKFPHLAVGEF 418

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1833	86.3	417	1 US-08-553-888A-3	Sequence 3, Appli
2	98	4.6	462	2 US-08-898-976-2	Sequence 2, Appli
3	98	4.6	462	2 US-08-898-976-4	Sequence 4, Appli
4	93.5	4.4	746	2 US-08-785-431-4	Sequence 4, Appli
5	93.5	4.4	788	2 US-08-785-431-2	Sequence 2, Appli
6	90.5	4.3	1248	2 US-08-726-214-16	Sequence 16, Appli
7	90.5	4.3	509	2 US-09-031-392-6	Sequence 6, Appli
8	90.5	4.3	509	4 US-09-299-549-6	Sequence 6, Appli
9	90	4.2	574	4 US-09-383-586-36	Sequence 36, Appli
10	90	4.2	759	4 US-08-637-759B-89	Sequence 89, Appli
11	90	4.2	759	3 US-08-871-355A-89	Sequence 89, Appli
12	88	4.1	357	1 US-08-119-773-2	Sequence 2, Appli
13	87	4.1	357	1 US-08-119-773-4	Sequence 4, Appli
14	87	4.1	357	1 US-08-119-773-6	Sequence 6, Appli
15	85.5	4.0	633	4 US-09-097-889-23	Sequence 23, Appli
16	85	4.0	336	3 US-08-749-816-2	Sequence 6, Appli
17	85	4.0	457	2 US-08-882-704A-6	Sequence 2, Appli
18	84.5	4.0	348	1 US-08-176-126B-2	Sequence 2, Appli
19	84.5	4.0	348	1 US-08-669-435-2	Sequence 2, Appli
20	84.5	4.0	348	5 PCT-US94-14431A-2	Sequence 2, Appli
21	84.5	4.0	473	1 US-08-597-236-13	Sequence 13, Appli
22	84.5	4.0	473	1 US-08-746-682A-13	Sequence 13, Appli
23	84	4.0	311	2 US-08-794-216-1	Sequence 1, Appli
24	84	4.0	429	2 US-08-677-049-5	Sequence 5, Appli
25	83.5	3.9	503	1 US-08-464-840-4	Sequence 4, Appli
26	83.5	3.9	503	1 US-08-483-094-4	Sequence 4, Appli
27	83.5	3.9	648	4 US-08-800-291B-8	Sequence 8, Appli

28	82.5	3.9	397	1 US-08-476-000-63	Sequence 63, Appli
29	82.5	3.9	397	1 US-08-472-840-63	Sequence 63, Appli
30	82.5	3.9	397	2 US-08-476-976-63	Sequence 63, Appli
31	82.5	3.9	397	3 US-08-474-410-63	Sequence 15, Appli
32	82	3.9	462	3 US-08-788-231A-15	Sequence 15, Appli
33	82	3.9	588	2 US-09-005-232A-2	Sequence 2, Appli
34	81.5	3.8	525	1 US-08-356-340-2	Sequence 2, Appli
35	81.5	3.8	525	2 US-08-786-555-2	Sequence 2, Appli
36	81	3.8	1285	2 US-08-540-406-6	Sequence 6, Appli
37	81	3.8	1285	4 US-08-656-055-6	Sequence 6, Appli
38	81	3.8	1285	4 US-08-954-668-6	Sequence 6, Appli
39	81	3.8	1285	5 PCT-US95-13233-6	Sequence 6, Appli
40	81	3.8	1286	4 US-09-268-140-3	Sequence 6, Appli
41	80.5	3.8	607	4 US-08-969-683A-59	Sequence 59, Appli
42	80.5	3.8	764	4 US-08-424-567-2	Sequence 2, Appli
43	80.5	3.8	764	4 US-08-711-928-2	Sequence 2, Appli
44	80.5	3.8	764	4 US-09-184-937-2	Sequence 2, Appli
45	80	3.8	999	2 US-08-473-533A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-553-888A-3
; Sequence 3, Application US/08553888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-553-888A-3

Query Match 86.3%; Score 1833; DB 1; Length 417;
Best Local Similarity 88.7%; Pred. No. 9.8e-185;
Matches 370; Conservative 12; Mismatches 35; Indels 0; Gaps 0;

QY 1 MSCXPRSVGRCLPLCALTLLEAALILLFYPTTHYDASLEDDKGLVASYOGQDTLWAAI 60
DB 1 MSSXPRSVGRCLPLCALTLLEAALILLFYPTTHYDASLEDDKGLVASYOGQDTLWAAI 60
QY 61 GLGFLTSSFRHRSSWSVAFLNLFMALGYVMAILLDGLFLSOPSGKVVITLFSIWLATWSA 120

Db	61	GLGFLTSSFRHSMSSVAFNLFMLALGQVNAILLDGLSOPRECKVITYLEFSTRILAMSA	120
Qy	121	LSVLSVDAYLGKYNLAOLVVMVLEVTdJGNTLRMTISNIFNEDYHNNMMHIYFAAYFG	180
Db	121	MSVLISAGAVLGKYNLAOLVVMVLEVTALGTLEMTVISNIFNEDYHNNLHEHVFYAAYFG	180
Qy	181	LTVMACLPKPLPECTEBDDQATIPISLSAMIGALFLMMF-PSVSAALLRSPIERKANFN	240
Db	181	LTVMACLPKPLPKGTEDDDQATIPISLSAMIGALFLMMFSPVSAALLRSPIERKANFN	240
Qy	241	TYVAVASVVAIAGSSSLAHPOGRISKTYGHSALVPEGVAVdTSCHLIPSPMLPIYGLGV	300
Db	241	TYVALAVSVVAIAGSSSLAHPOGRISMTYVHASVAVLAGVAVGTSCHLIPSPMLAVGLGV	300
Qy	301	AGLISVGAATLPCCCNRVLCIPHSSTMGVNFSLGLLEETIYVLLVLDTVGANGMIG	360
Db	301	AGLISIGCAKCLPVCNCRVLCIHHISVMHSIFSLGLIGETIYVLLVLTWVNGNMIG	360
Qy	361	FQVLLSTIGELSLAVIALTSLGLTALLNLKIFAPHEAVYEPDQVFWKRPHLAVGF	417
Db	361	FQVLLSTIGELSLAVIALTSLGLLLNLKIMAPHVAVYFDQVFWKRPHLAVGF	417

```

1      RESULT      2
2      US-08-898-976-2
3      Sequence 2, Application US/08898976
4      Patent No. 5891670
5      GENERAL INFORMATION:
6      APPLICANT: Burnham, Martin
7      APPLICANT: Lonetto, Michel
8      APPLICANT: Warren, Patrick
9      TITLE OF INVENTION: NOVEL TETRACYCLINE RESISTANCE
10     TITLE OF INVENTION: PROTEIN
11     NUMBER OF SEQUENCES: 6
12     CORRESPONDENCE ADDRESS:
13     ADDRESSEE: dechert Price & Rhoads
14     STREET: 997 lexott Drive, Building 3, Suite 210
15     CITY: Lawrenceville
16     STATE: NJ
17     COUNTRY: USA
18     ZIP: 08543
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Diskette
21     COMPUTER: IBM Compatible
22     OPERATING SYSTEM: DOS
23     SOFTWARE: Fastseq for Windows Version 2.0
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/898,976
26     FILING DATE:
27     CLASSIFICATION: 435
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER:
30     FILING DATE:
31     ATTORNEY/AGENT INFORMATION:
32     NAME: Bloom, Allen
33     REGISTRATION NUMBER: 29,135
34     REFERENCE/DOCKET NUMBER: GM10044
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: 609-520-3214
37     TELEFAX: 609-520-3259
38     TELEX:
39     INFORMATION FOR SEQ ID NO: 2:
40     SEQUENCE CHARACTERISTICS:
41     LENGTH: 462 amino acids
42     TYPE: amino acid
43     STRANDEDNESS: single
44     TOPOLOGY: linear
45     US-08-898-976-2
46
47     Query Match      4.68; Score 98; DB 2; Length 462
48     Best Local Similarity 20.08; Pred. No. 0.065;

```

	Matches	96;	Conservative	70;	Mismatches	162;	Indels	152;	Gaps	21.
QY	1	MSCKIPRSVGRCLP	CACTLTLEAAILLFFPHNYASLEDDGGLVASIYVGDDLYVMAAI	60						
Db	13	MNVEYSK-IKRAVPI	-----LLEFLFVSLINDSFR-----LIVAIADLDNI----	54						
QY	61	GLGFLTSSFRRRSMESVAFNFMALG	QOMAILLDGFLSQ--FPSPGVYTLTFSLATM	118						
Db	55	-----SVTVYSQKATLAGLVIGMGAVVYASLS	DAISRPFYIGVILIIIGSIIIGNTF	107						
QY	119	SALSVLISVDVAVLGNVLAOLVVMVLEV	-----TdlG-----NLRMYISN---	159						
Db	108	QHOSPLTLVGRILIQAGLAAETLVIYAKYLSKEDQTYIGLSTSYSLSVLIGTSLG	167							
QY	160	---IFPTDYNHNMNMHYVPAAYFGLVAMCIP---	KPLREGTEIDQCRATIPSLSAMLGL	214						
Db	168	GFITFLYHMTNMLI-----ALIVYETLPFLFLRK--	ENNNKKAHLDDVGLLIVAT	218						
QY	215	-----FLMFRPVSNSALL-----	RSPIERKNAVENTYVAVASVTAIS	254						
Db	219	IATTVMLFITPENNMLYGALIAITFAFYALIKNAQRPLVNRKSFONKRASFLFIVMY	278							
QY	255	GSSLAHPOGRKISKTYGHSAVLP-----	EGVAVG--TSCHLIPSPMLPVLGLVAGLISVR	307						
Db	279	AIQL-----GYLTFPPIPEQIYHIDLDTTSLSLVGLGYVAAYVGLSGAT---	324							
QY	308	GAKYLPGCCNREVLGIPHSIMGYNSLGLLEELIYIYL--	LVLDDVGAANGAIGEVLL	365						
Db	325	-GEYLN-----SKQAIIPAILIALSLILPFAVGNHISIFVISMI	364							
QY	366	-----SIGLSLAIYALSLGL--LWALLNMLKIRKP	396							
Db	365	FPAGSFALMYPDLNLEAIKTLDLNTGVAIGYLLINVAASVGAIAAALDFALANP	424							

RESULT 3
US-08-898-976-4
Sequence 4, Application US/08898976
Patent No. 5891670
GENERAL INFORMATION:
APPLICANT: Burnham, Martin
APPLICANT: Lonetto, Michael
APPLICANT: Warren, Patrick
TITLE OF INVENTION: NOVEL TETRACYCLINE RESISTANCE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
City: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/08/898, 976
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: GM10044
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 462 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-898-976-4

```

Query Match 4.6%; Score 98; DB 2; Length 462;

Best Local Similarity 20.0%; Pred. No. 0.065; Mismatches 96; Conservative 70; Mismatches 162; Indels 152; Gaps 21;

```

QY 1 MSCKPRSVGKCLPLCALTEALILFEFFTHYDASLEDOGLVAVGODLVYMAAI 60
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 13 MNVEYSK-IKKAVPI-----LLFLFVSLVIDNSFK-----LISVALADLNI----- 54
QY 61 GLGFLTSSFRHRSWSVAFLFMLALGVOMAILDGLSQ--PPSGKAVITLESWLATM 118
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 55 -----SVTVSWQATLAGLVIGMGAVVASLSDAISIRPFYGVILLIFGSIIGNIF 107
QY 119 SALSVLSVAVLGKVMLAQVVMVLEVE-----TDLG-----NLRMTISN--- 159
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 108 QHOSPLLVGRITOTAGLAAETLYIVYAKYLSKEDQKTYLGLSTSSYSLSIVTISG 167
QY 160 --IFNDYHNMNMHIYVFAAYFGLTVAWCLP--KPLPEGETEDDO:ATIPSLSAML 214
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 168 GFITVYLMHNMELI-----ALIVFTLPEFLKLPK--ENNTKNNAHLDVGLLIVAT 218
QY 215 -----FLMFTPPSANSALL-----RSPIERKNAVFNTYVAVAVVTAIS 254
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 219 IATVMLEFTNFMLVIGALIAIIVFALYIKNAQRPLVNSFPQNKRYASFLEIVFEM 278
QY 255 GSSLAHQGRKSKYGHSAVLP-----EGYAVG-TSCHLIPSPWLDIVGLVAGLISV 307
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 279 AIQV-----GYTFPFTEQIYHLQDFTSLIPGVYIVAVIGALSGLK1-- 324
QY 308 GAKYLPCCNVLGIPSHSSINGYNSLLGLEETIYVL-LVLDVYAGNMGIFOVLL 365
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 325 -GELTN-----SKQAITAILIALSLILPAFVAGNHSIFSVISMT 364
QY 366 -----SIGELSLAIVIALTSG-LTALLNLKIRKAP 396
Db 365 FFAGSFALMAYPLINEAIFKTIIDLMNTGVAIGVFYLLINVAVSGLATAALIDPKALNEP 424

```

```

RESULT 4
US-08-785-431-4
; Sequence 4, Application US/08785431
; Patent No. 5891667
; GENERAL INFORMATION:
; APPLICANT: Chalke, Alison
; APPLICANT: Hodgson, John
; TITLE OF INVENTION: No. 5891667el Spoilite
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,431
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9600955.0

```

```

; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 746 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-431-4

```

Query Match 4.4%; Score 93.5; DB 2; Length 746;

Best Local Similarity 24.2%; Pred. No. 0.4; Mismatches 59; Conservative 37; Mismatches 87; Indels 61; Gaps 13;

```

QY 199 DQATIPSLSAMLALFLMFRPSVNSALLRSPIERKNVAF-NTYVAVAVSV-VTAISGS 256
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 268 EEAGEVTNVSYPVPLTL-LNQPAKOKATSKAEVQRGOVLENTLKDFGVAKVTOI--- 323
QY 257 SLAHQGRKISKTYGHSVLP--EGYAVGTSCHLIPSPWLDIVGLVAGLISVGA 315
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 324 -----KIGPAVTOYETLOPQGVKSKIVNLND---TALLAAK--DVRIEAPIPGR 370
QY 316 CNRYLGIPIHSSIMGNFSLGLEETIYVLLVD-----TVGAGNMGIFOVLLS 366
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 371 SAVGIEVFNENI-----SLVSLKE-----VDEKFPNSNKLVEGLGRISDPIITVP 417
QY 367 IGEISLAIVIALT-----SGLLTALLNLKIRKAPHEAKYF-----DQVFNKFP 411
   |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 418 LNEPHLLVAGSTGSGKVCINGIITSILLNAK---PHEVKMLIDPKKWEVLVWYNGIP 473
QY 412 HLAIV 415
Db 474 HLLI 477

```

```

RESULT 5
US-08-785-431-2
; Sequence 2, Application US/08785431
; Patent No. 5891667
; GENERAL INFORMATION:
; APPLICANT: Chalke, Alison
; APPLICANT: Hodgson, John
; TITLE OF INVENTION: No. 5891667el Spoilite
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,431
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9600955.0
; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gilm, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31349
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 788 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-785-431-2

Query Match 4.48; Score 93.5; DB 2; Length 788;
 Best Local Similarity 24.28; Pred. No. 0.44;
 Matches 59; Conservative 37; Mismatches 87; Indels 61; Gaps 13;

QY 199 DQRTIPSLMGLFLMFRPSVNSALRSPIERKNAVF-NYYVAASV-VTAISGS 256
 DB 310 EEAEEVNVSVVPEPLT-LMOPAKOKATSKAEVQRKQVLENTLKDFGVAKVYQI--- 365
 QY 257 SLAPROCKISKTYGSAVLP-EGVAVDTSCHLIPSPMLPIVLGLVAGLISYVGAKYLPQC 315
 DB 366 -----KIGPAVTVYEIQPQGVKSVKIVLHND---IALALAAK--DVRIEAPIPGR 412
 QY 316 CNRVLGIPHSSIMKYNLSLGLLEIYIVLVLD-----TVGAGNMICFOYLLS 366
 DB 413 SAVGIEVPEKRI-----SLVSLKE-----VLDEKFPSSNNKLEVGIRDISGDPITVP 459
 QY 367 IGEISLAIIVALT-----SGLTALLNLKIRKAPHEAKYF-----DDQVFMKFP 411
 DB 460 LNEPHLVAGSTSGSKSVCINGIITLILNAK-----PHEVKMLIDPKMVELNVINGIP 515
 QY 412 HLAV 415
 DB 516 HLLI 519

RESULT 6
 US-08-726-214-16
 Sequence 16, Application US/08726214
 Patent No. 6107076
 GENERAL INFORMATION:
 APPLICANT: Tang, Wei-Jen
 APPLICANT: Gilman, Alfred G.
 TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
 TITLE OF INVENTION: AND USES THEREFOR
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: United States of America
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/726,214
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION NUMBER:
 APPLICATION NUMBER: US 60/005,498
 FILING DATE: 04-OCT-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: UTSD:450
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (512) 474-7577
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1248 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-726-214-16.

Query Match 4.38; Score 91; DB 3; Length 1248;
 Best Local Similarity 23.08; Pred. No. 1.6;
 Matches 45; Conservative 40; Mismatches 87; Indels 24; Gaps 6;

QY 24 LILFEFTHYDASLEDOKGLVASVYQVODLTWMAAIGLFTSSFRHSMS---VAF 79
 DB 191 LTLVLHLASAPMDPLKGLIGFTGIEVICALVVRKDTSHYLYSGVVTWVM 250
 QY 80 NLEMLAGVQMAIILDGFLSQPSGKVYITLFSIWLATMSALSVLSVDVLCVNIQL 139
 DB 251 TQIILAAGLGVLGDGI-----GYVLFLLFARY--SMPLPLTWAILLGLG-TSLIY 301
 QY 140 VVMVLVEVTDIGNLWVINSNFTDYMNNMHYVFAAYG-----LTVMACLPKL 191
 DB 302 TLQYLIRLAVFSINQVLAQV---LPMCMNTAGIFISYSDRAQRAFLETRCYEARL 358
 QY 192 PEGTEDNDQATIPSL 207
 DB 359 RLETFENQROERLVLSV 374

RESULT 7
 US-09-031-392-6
 Sequence 6, Application US/09031392
 Patent No. 5942398
 GENERAL INFORMATION:
 APPLICANT: Tartaglia, Louis A.
 APPLICANT: Weng, Xun
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkielejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 TYPE: amino acid
 TOPOLOGY: linear


```

APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-637-759B-89

```

```

Query Match          4.2%; Score 90; DB 2; Length 759;
Best Local Similarity 20.5%; Pred. No. 0.97;
Matches 75; Conservative 59; Mismatches 125; Indels 106; Gaps 17;

QY 52 QDLTVMAAIGLGLTSSFRHWS-----SVAFLMLALGVQAILLDGFLSQFPG 104
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 QDMVLATVLLAIYVMLLPCTWVDILITINIMFSVILLIAIYLSDDPD--LSVFP 157
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 KVVITLFSIATMSALSVALS-----VDA-----VLGKYNLAQVYVWVLEVDGLNLR 154
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LLITTLRLSLTSTSRVLLOHAGNIVDAFGKVVGGMLTVGLVFTTITVQF----- 213
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 MVISNIENFDYHMMMHIIYFAVFGLT-----VAMCLPKPLPGTEDDQFATIPS----- 206
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 IVITK-----GIEVAEVSARFSLDGMFGKQMSIDGLRAGVIDADHARTLRQHVQ 265
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 -----LSAMLGALFLMMPSPVNSALLSPIERKNAVFNTYAAVSVYTAISGSSLAHPQ 262
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 ESRLFGAMDGM-----KFKVGDITAGIIVLVNIIIGGIITIAIVQ 305
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 263 GKIS-----KTYGSAVLPEGVAVDTSCHLIPSPWLPVIGLVAGLSVCGAKYLPG--- 314
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 YDMSMSAVHTY---SVLSIG---DGLCGQIPS---LLISLSAGIIVTR---VPEEKR 351
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 315 -----CCNRYLGIPIHSSIM-----GYNFSLLGLAEIIVYVLLVLT---DTVG 353
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 ONLATELSSQIARQPSQLITLAVVLMALLIPGPFITLAFFSALLPILIRKKS SV 411
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 AGNGM 358
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 SANCV 416

```

```

RESULT 11
US-08-871-355A-89
Sequence 89, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-871-355A-89

```

```

Query Match          4.2%; Score 90; DB 3; Length 759;
Best Local Similarity 20.5%; Pred. No. 0.97;
Matches 75; Conservative 59; Mismatches 125; Indels 106; Gaps 17;

QY 52 QDLTVMAAIGLGLTSSFRHWS-----SVAFLMLALGVQAILLDGFLSQFPG 104
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 QDMVLATVLLAIYVMLLPCTWVDILITINIMFSVILLIAIYLSDDPD--LSVFP 157
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 KVVITLFSIATMSALSVALS-----VDA-----VLGKYNLAQVYVWVLEVDGLNLR 154
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LLITTLRLSLTSTSRVLLOHAGNIVDAFGKVVGGMLTVGLVFTTITVQF----- 213
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 MVISNIENFDYHMMMHIIYFAVFGLT-----VAMCLPKPLPGTEDDQFATIPS----- 206
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 IVITK-----GIEVAEVSARFSLDGMFGKQMSIDGLRAGVIDADHARTLRQHVQ 265
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 -----LSAMLGALFLMMPSPVNSALLSPIERKNAVFNTYAAVSVYTAISGSSLAHPQ 262
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 266 ESRLFGAMDGM-----KFKVGDITAGIIVLVNIIIGGIITIAIVQ 305
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 263 GKIS-----KTYGSAVLPEGVAVDTSCHLIPSPWLPVIGLVAGLSVCGAKYLPG--- 314
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 YDMSMSAVHTY---SVLSIG---DGLCGQIPS---LLISLSAGIIVTR---VPEEKR 351
   ||: : : : : : : : : : : : : : : : : : : : : : : : : : :

```

OY 315 -----CCNRVLGIPHSIM-----GYNFSLGLGELIYIVLVL---DTVG 353
 Db 352 QNATELSQIARQPOSILITAVVLMALIPGPFITLAFPSALALPILIRKRSV 411
 OY 354 AGNGM 358
 Db 412 SANGV 416

RESULT 12

US-08-119-773-2

; Sequence 2, Application US/08119773

; Patent No. 5460942

; GENERAL INFORMATION:

; APPLICANT: Chou, Janice Y.

; APPLICANT: Shelly, Leslie L.

; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND

; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Hourie and Crew

; STREET: Stewart Street Tower, One Market Plaza

; CITY: San Francisco

; STATE: California

; COUNTRY: US

; ZIP: 94105-1493

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/119,773

; FILING DATE: 10-SEP-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Kenneth A.

; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 15280-175

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 357 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-119-773-2

Query Match 4.18; Score 88; DB 1; Length 357;

Best Local Similarity 19.98; Pred. No. 0.5;

Matches 75; Conservative 52; Mismatches 116; Indels 134; Gaps 18;

OY 20 LEALILILFFFTHYDASLEDOKGLVASVOYGDITVMAAIGLGLTSSFRHSSVAF 79
 Db 39 LRNFYVLEFPIWHLQEA-----VGILMWAVIG-----DMLNLVF 75
 OY 80 NLEMLALGVOMAILLDGFLS-----QF-----PSGRVITLFSIWTMSAL 121
 Db 76 KWILFGORPYWVLDITVYTSVPLIKQFPVTCETGPGSPSGHAMGTAGYVWVSTL 135
 OY 122 SYL-----ISVDVGLKVNLA-----QVYVNVL--VEYTLG 151
 Db 136 SIFGKIKPIYRERCLNVILMLGFWAQLNVCLSRIYLAHFPQVAVGVLGAVETTF 195
 OY 152 NLRVSIENFTD---YHNMNMHIYFAAYF-----GLVAVCLPK-----PLPGE 196
 Db 196 S---HHSIVASAKKYKFLIFLFFLFSRAIGYLLKGLGVLLMTLEAKQWCEQPEVNH 252
 OY 197 DNDQATIP--SLSAMIGALF-LMNF-PSVNSALLRSPIERKNVENVTYAVAVSVYTAI 253

Db 253 ID-----TFPFASLKNLGTLEGLL---ALNMSYRSC----- 284
 OY 254 SGGSLAHPOGKISK--TYGSAVILDEGVAVGTSTCHLIPSPMLPIVLGV-----AGLISY 307
 Db 285 -----KGKLSKWLPRLSIVASLVLLHVPDSLRPPSOVELFVYLSCKSAVYPLA 336
 OY 308 GAKYLPGCCNRVLGIPH 324
 Db 337 SVSVIPLCLAQVLGPH 353

RESULT 13

US-08-119-773-4

; Sequence 4, Application US/08119773

; Patent No. 5460942

; GENERAL INFORMATION:

; APPLICANT: Chou, Janice Y.

; APPLICANT: Shelly, Leslie L.

; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND

; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Hourie and Crew

; STREET: Stewart Street Tower, One Market Plaza

; CITY: San Francisco

; STATE: California

; COUNTRY: US

; ZIP: 94105-1493

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/119,773

; FILING DATE: 10-SEP-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Kenneth A.

; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 15280-175

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 357 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1..357

; OTHER INFORMATION: /label= R to C at 83

; OTHER INFORMATION: /note= "The G-6-Pase amino acid sequence mutated

; OTHER INFORMATION: from Arg to Cys at position 83. "

; US-08-119-773-4

Query Match 4.18; Score 87; DB 1; Length 357;

Best Local Similarity 19.98; Pred. No. 0.63;

Matches 75; Conservative 52; Mismatches 116; Indels 134; Gaps 18;

OY 20 LEALILILFFFTHYDASLEDOKGLVASVOYGDITVMAAIGLGLTSSFRHSSVAF 79
 Db 39 LRNFYVLEFPIWHLQEA-----VGILMWAVIG-----DMLNLVF 75

```

QY 80 NLFMALGVOMAILLDGFLS-----QF-----PSGKVITLPSIWLATMSAL 121
;      :      :      :      :      :      :      :      :      :
Db 76 KWILFGQCPYMWVLDTDYTSVPLIKQFPVTCETGPGSGSHAMGTAGYIYMWSTL 135
;      :      :      :      :      :      :      :      :      :
QY 122 SVL-----ISVDAVLGKVNL-----QLYVMVL--VEVTDLG 151
;      :      :      :      :      :      :      :      :      :
Db 136 SIFGKIKPTVRFRCNLVILMLGFWAVOLNCLSRILYLAHFPHQVAVAGVLSGIATVET 195
;      :      :      :      :      :      :      :      :      :
QY 152 NLRVINSIFNTD---YHNMNMHIYVPAVF-----GLTVANCLPK-----PLPGTE 196
;      :      :      :      :      :      :      :      :      :
Db 196 S---HHSIYNASLKKYFLITFFLFSFRAIGFYLLKGLGVDLMTLEKAQRWCQEPV 252
;      :      :      :      :      :      :      :      :      :
QY 197 DNDQATIP--SLSAMLGALF-LMMFPPSVNSALLRSPERKNNAVFNTYAVAVSVTAI 253
;      :      :      :      :      :      :      :      :      :
Db 253 ID-----TTPFASLKNIGTLFGGL--ALNSSMYRESC-----284
;      :      :      :      :      :      :      :      :      :
QY 254 SGSSLAHPQCKISK--TYGHSNAVPEGVAVATSCHLIPSPMLPIVLGLV---AGLISvr 307
;      :      :      :      :      :      :      :      :      :
Db 285 -----KGLSKWLPFLCSSLIVASLVLLHVPDSLKPSPQVELVYFVLSFCKSAVVPLA 336
;      :      :      :      :      :      :      :      :      :
QY 308 GAKYLPCCNRYLGIPH 324
;      :      :      :      :      :      :      :      :      :
Db 337 SVSVIPYCLAQVLGQPH 353
;      :      :      :      :      :      :      :      :      :

RESULT 14
US-08-119-773-6
; Sequence 6, Application US/08119773
; Patent No. 5460942
; GENERAL INFORMATION:
; APPLICANT: Chou, Janice Y.
; APPLICANT: Lei, Ke-jian
; APPLICANT: Shelly, Leslie L.
; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND
; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,773
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..357
; OTHER INFORMATION: /label- R to C at 295

```

```

; OTHER INFORMATION: /note- "G-6-Pase Mutated at position 295 with Arg
; OTHER INFORMATION: changed to Cys"
US-08-119-773-6

Query Match      4.18; Score 87; DB 1; Length 357;
Best Local Similarity 19.98; Pred. No. 0.63;
Matches 75; Conservative 52; Mismatches 116; Indels 134; Gaps 18;

QY 20 LEALLILFFFTHYDASLEQOKLVASYOGDLYMAAIGLFLSSPFRHSSVAF 79
;      :      :      :      :      :      :      :      :      :
Db 39 LRNAFYVLFPMFLQEA-----VGKLLMWAVYG-----DMLNLF 75
;      :      :      :      :      :      :      :      :      :
QY 80 NLFMALGVOMAILLDGFLS-----QF-----PSGKVITLPSIWLATMSAL 121
;      :      :      :      :      :      :      :      :      :
Db 76 KWILFGQCPYMWVLDTDYTSVPLIKQFPVTCETGPGSGSHAMGTAGYIYMWSTL 135
;      :      :      :      :      :      :      :      :      :
QY 122 SVL-----ISVDAVLGKVNL-----QLYVMVL--VEVTDLG 151
;      :      :      :      :      :      :      :      :      :
Db 136 SIFGKIKPTVRFRCNLVILMLGFWAVOLNCLSRILYLAHFPHQVAVAGVLSGIATVET 195
;      :      :      :      :      :      :      :      :      :
QY 152 NLRVINSIFNTD---YHNMNMHIYVPAVF-----GLTVANCLPK-----PLPGTE 196
;      :      :      :      :      :      :      :      :      :
Db 196 S---HHSIYNASLKKYFLITFFLFSFRAIGFYLLKGLGVDLMTLEKAQRWCQEPV 252
;      :      :      :      :      :      :      :      :      :
QY 197 DNDQATIP--SLSAMLGALF-LMMFPPSVNSALLRSPERKNNAVFNTYAVAVSVTAI 253
;      :      :      :      :      :      :      :      :      :
Db 253 ID-----TTPFASLKNIGTLFGGL--ALNSSMYRESC-----284
;      :      :      :      :      :      :      :      :      :
QY 254 SGSSLAHPQCKISK--TYGHSNAVPEGVAVATSCHLIPSPMLPIVLGLV---AGLISvr 307
;      :      :      :      :      :      :      :      :      :
Db 285 -----KGLSKWLPFLCSSLIVASLVLLHVPDSLKPSPQVELVYFVLSFCKSAVVPLA 336
;      :      :      :      :      :      :      :      :      :
QY 308 GAKYLPCCNRYLGIPH 324
;      :      :      :      :      :      :      :      :      :
Db 337 SVSVIPYCLAQVLGQPH 353
;      :      :      :      :      :      :      :      :      :

RESULT 15
US-09-097-889-23
; Sequence 23, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrnsadt, Corina
; APPLICANT: Ghosh, Soumitra S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

```



```

; INFORMATION FOR SEQ ID NO: 23
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 603 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
US-09-097-889-23

```

Query Match	4.0%;	Score 85.5;	DB 4;	Length 603;
Best Local Similarity	17.4%;	Pred. No. 2;		
Matches 95;	Conservative 79;	Mismatches 177;	Indels 195;	Gaps 21;

```

0Y 5 YPRSVGRCLPCLALEAAILLEAAILLYFTTHNDASLEEDOKGVASOVQODTLVAAIGLE 64
Db 32 YPHYK-----SIAVSTIISLEPTTMECL-DOEVIISNMHATTOFTO--LSLF 80
0Y 65 LTSSFRH-----SNSSVAFNLFMLALGVOWAAILDLDFLSQFSGKAVITLFSIWA 116
Db 81 KLDFESMFIPLVAFLEVTWSIMEFSL-----WYMSDPNINQFEKYLIIPLITMILIV 132
0Y 117 TMSAL-----SVLISVDVAVLGKVLNLAQLVNAVLYEVTDGCLRNVISNIEN 162
Db 133 TANNLFOLFIEGEGVINSPELLISMWYARADANTAALQAILVYNIGDIG----- 181
0Y 163 TDYHNNMHIIYFAAYGGLVAVMCLPRPREGTEDDIDRATF--PSLSMGLAFL-- 216
Db 182 -----FLTALAMFT-----LHSNSWDPQOMALLNANPSLTPLGLLTLAAAG 222
0Y 217 -----W--MFPVSVALRS-----PIERKNAVFNTRYAVAS 248
Db 223 KSAQGLGHPMLPSAMEGTTPYALSLLHSTWVAGIFILLIFHPLAESPLIOT-LTLCG 281
0Y 249 VVTALSGSSLAHPQCKISKTYGHSAVLpgevavdT-----SCHLPSWLPjVL- 297
Db 282 AITTFEAVCALTONDIKKIVAFSTSSQLGIMVYIGINOPHLAETHICTHAEFKAWLFEM 341
0Y 298 -----GVAGLSLV-GAKYLPBCCCR-----VLGI 322
Db 342 CSGSIHNLNNEODIRKMGGLTKMPLTSLTSLTGSLAGMPLTGFYSKDHIIETAM 401
0Y 323 PHSISMGVFSLLGLAEIIVYLVEDT-----VGAGN 356
Db 402 SYTNMMAISITLINTSLTSAVSTGMILITLGGORPFTPLINENNPTLNPITKRLAAGS 461
0Y 357 GMIGFOVLLSIG-----ELSLAIVALT-----GLTLALLN-----LKIRAPHEAK 400
Db 462 LFAGELLNINISPAAPFQTTIPLVYKLTALAVTFLGTLTALDLNLYTNKLM-KSPLCF 520
0Y 401 YFDQV 406
Db 521 YFSNML 526

```

THIS PAGE BLANK (USPTO)

OM of: US-09-600-714-41 to: Issued_Patents_AA.* out_format : pfs

Date: Sep 11, 2001 8:40 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE1=frame+ncp.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09600714/runat.11092001.083823.22058/app.query.fasta.1.1332
-DB=Issued_Patents_AA -QFMT=fasta -SUFFIX=rai -CAP=12.000
-CAPERT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-GAPDP=4.500 -GAPEXT=0.050 -XGAP=10.000 -XGAPEXT=0.500
-FCAPDP=6.000 -FCAPEXT=7.000 -YCAP=10.000 -YCAPEXT=0.500
-DELDP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pcr
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORTEXT -MINLEN=0 -MAXLEN=200000000
-USER=US09600714 -CGN1_1_20 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-09-600-714-41
Query length: 1254
Database: Issued_Patents_AA.*
Database sequences: 197339
Database length: 20590346
Search time (sec): 60.250000

score_list:

Sequence	Strd Orig	zscore	EScore	len	Documentation	..7e-187	417
/cgn2.6/prodata/2/1aa/5A.COMB.pcp:US-08-553-888A-3			1934.00	3484.43	2.7e-187	417	1
/cgn2.6/prodata/2/1aa/5B.COMB.pcp:US-08-483-533-41			124.00	198.99	0.0003	355	1
/cgn2.6/prodata/2/1aa/5C.COMB.pcp:US-08-635-121-2			110.50	158.65	0.0149	1346	1
/cgn2.6/prodata/2/1aa/5D.COMB.pcp:US-08-933-803A-16			108.50	163.28	0.0165	671	1
/cgn2.6/prodata/2/1aa/5E.COMB.pcp:US-08-726-214-16			106.50	152.28	0.0364	1248	1
/cgn2.6/prodata/2/1aa/5F.COMB.pcp:US-08-149-097D-35			105.50	143.99	0.0524	2509	1
/cgn2.6/prodata/2/1aa/5G.COMB.pcp:US-08-466-033-183			105.00	162.08	0.0276	469	1
/cgn2.6/prodata/2/1aa/5H.COMB.pcp:US-08-444-733-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5I.COMB.pcp:US-08-464-134-15			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5J.COMB.pcp:US-08-461-361-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5K.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5L.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5M.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5N.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5O.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5P.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Q.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5R.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5S.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5T.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5U.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5V.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5W.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5X.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Y.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Z.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5A.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5B.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5C.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5D.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5E.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5F.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5G.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5H.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5I.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5J.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5K.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5L.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5M.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5N.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5O.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5P.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Q.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5R.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5S.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5T.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5U.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5V.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5W.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5X.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Y.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Z.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5A.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5B.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5C.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5D.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5E.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5F.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5G.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5H.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5I.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5J.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5K.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5L.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5M.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5N.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5O.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5P.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Q.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5R.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5S.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5T.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5U.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5V.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5W.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5X.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Y.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Z.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5A.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5B.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5C.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5D.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5E.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5F.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5G.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5H.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5I.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5J.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5K.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5L.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5M.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5N.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5O.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5P.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Q.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5R.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5S.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5T.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5U.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5V.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5W.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5X.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Y.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5Z.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5A.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5B.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5C.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5D.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5E.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5F.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5G.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/1aa/5H.COMB.pcp:US-08-485-910-183			105.00	139.51	0.0803	2910	1
/cgn2.6/prodata/2/							

```

201 GAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTTCATGC 250
   |:::|
67 rAsnheArghisSerTIPserSerValAlaIphAsnLeuPheMetL 84
251 TGGCGCTGTGTGCGAGTGGGCAATCCTGTGAGCGGCTCTCGAGCCAG 300
   euAlaLeuGlyValIGInTPraIaIleuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTGGGAAGTGTGATCATCACTGTTCAGTATTGGCTGGCCAC 350
   |:::|
101 PheProPGIlySValValIleThrLeuPheSerIleArgLeuAlaIth 117
351 CATGAGCTCTTGTGCTGATCATCTGGATGCTGTCTTGGGGAGG 400
   |:::|
117 rMetSerIaMetSerValIleuIleSerAlaGlyAlaValLeuGlySv 134
401 TCACCTTGGCGAGTGTGTGATGTGTGTGAGGTGAGGTGACAGCTTGA 450
   |:::|
134 aIAsnLeuAlaGInLeuValValMetValLeuValGInValThraIaLeu 150
451 GGCAACCTGAGATGTGATCATGATATCTTCAACACAGACTACCAT 500
   |:::|
151 GlyThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTTCAGACCTATTGGCGTGTGTGG 550
   |:::|
167 tAsnLeuArgHisPheTyrValIlePheAlaIaIatyrPheGlyLeuThrVal 184
551 CCTGTGCTGCGCAAGCTCTACCCGAGGAGGAGGAGGATTAAGATCAG 600
   |:::|
184 lATrPCysLeuProIysProIeuProIysGlyThrGluAspAsnAspGln 200
601 ACAGCAAGATACCCAGTTGTCTGCGCATGCTGGCGCTCTCTTGTG 650
   |:::|
201 ArgAlaThrIleProSerLeuSerIaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTGGCCAACTTCAACTCTGCTGTCTGTGAGAGTCCATCGAAA 700
   |:::|
217 pMerThrTrProSerValAsnSerProIeuLeuArgSerProIleGln 234
701 GGAGAAATGCCGTGTTCACACACTACTGTGTGTGTGTGTGTGTGTG 750
   |:::|
234 rGlySAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCATCTCAGAGGTGATCCTGTGCTACCCGAGGAGGATGAGCAA 800
   |:::|
251 ThrAlaIleSerGlySerSerIleuAlaHisProGlnArgIysIleSerMe 267
801 GACTTATGTGACAGTGGGTGTGTGGACGAGGCGTGGCTGTGGTACT 850
   |:::|
267 tThrTyrValHisSerAlaValIleuAlaGlyGlyValAlaValGlyTrs 284
851 CGTGACACCTATCCCTCTCCGTGGCTGGCCATGAGTGTGTGTGTG 900
   |:::|
284 erCysHisLeuIleProSerProTrIleuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTAA 950
   |:::|
301 AlaGlyLeuIleSerIleGlyGlyAlaIaIysCysLeuProValCysCysAs 317
951 CCGAGTGTGGGATTCGCCACAGCTCCATGAGGCTTACAACTTCACACT 1000
   |:::|
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGTGTGGTCTGCTGTGGAGAGATCATCTACATTTGTCTGTGTGTGTAT 1050
   |:::|
334 euLeuGlyLeuLeuGlyGInIleThrTyrIleValIleuLeuValLeuHis 350
1051 ACCGTGCGAGCGCGCATGATGATGCTTCCAGGTCCTCCAGCAT 1100
   |:::|
351 ThrValITrPAsnGlyAsnGlyMetIleGlyPheGlnValIleuLeuSerIl 367

```

```

1101 TGGGAACCTCAAGTCTGGCCATGCTGATAGTCTCTCAGCTGTGCTCTGA 1150
   |:::|
367 eGlyIuLeuSerLeuAlaIleValIleAlaIleuThrSerIlyLeuLeuT 384
1151 CAGTTTGTCTCTTAATCTTAATATGGAAGCAACCTCATGAGGCTTAA 1200
   hrcGlyLeuLeuAsnLeuAsnLeuIysIleTrpIysAlaProHisValAlaIaIys 400
1201 TATTTTGTATGACCAAGTTTCTGTGGAAGTTTCTCATTTGGCTGTGAT 1250
   |:::|
401 TyrPheAspAspGInValIleThrTrpIysPheProHisLeuAlaValGlyPh 417
1251 T 1251
   |
417 e 417

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-483-533-41
seq_documentation_block:
; Sequence 41, Application US/08483533
; Patent No. 6172047
;
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; TITLE OF INVENTION: Diseases
; NUMBER OF SEQUENCES: 43
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
;
; PRIOR APPLICATION DATA: 07/861,233
; APPLICATION NUMBER: 31-MAR-92
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
;
; TELEX: 25-3856
;
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-483-533-41

alignment_scores:
  Quality: 124.00      Length: 312
  Ratio: 1.033        Gaps: 19
  Percent Similarity: 38.462      Percent Identity: 24.679

alignment_block:

```

US-09-600-714-41/rev x US-08-483-533-41 ..

Align seg 1/1 to: US-08-483-533-41 from: 1 to: 355

```

921 CCCGAGGAGATCAGCCGACAGCAGACCCAGCACCATTGCGACGCCAG 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 ProthrglyAlaValProthrlaGlInserGlnValThrSerThrProAs 34
    : : : : : : : : : : : : : : : : : : : : : : : : :
871 GAGAGGATCAGGTGACACGAGGTACCCAGCAGCAGCAGCCTCTGCAAC 822
    : : : : : : : : : : : : : : : : : : : : : : : : :
34 nserGlnProAlaVal.ArgSerAlaProAlaAlaAlaProPro.ProPr 50
    : : : : : : : : : : : : : : : : : : : : : : : : :
821 ACCGCACTGTGCACATTAAGTCTTGCTGATCTTCCTTGGGGGTAGCCA. 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 oPro.....AlaGlyGlyProProp 57
    : : : : : : : : : : : : : : : : : : : : : : : : :
772 .....AGATGACCCCTGAGA...TGGCTGTACACCGCTGACTGCTA 734
    : : : : : : : : : : : : : : : : : : : : : : : : :
57 roSerGlySerLeuLeuLeuArgGlnTrpLeu.His..... 68
    : : : : : : : : : : : : : : : : : : : : : : : : :
733 CAGCATAGTAGGTGTGAACAGCGCATCTTCCTTGCATTGAGACTTCTC 684
    : : : : : : : : : : : : : : : : : : : : : : : : :
69 .....ValPro...GluSerAlaSerA 75
    : : : : : : : : : : : : : : : : : : : : : : : : :
683 AGCAGAGCAGAGTTGAAACTTGGCCGACAGATCCACAGAAAGAGG... 638
    : : : : : : : : : : : : : : : : : : : : : : : : :
75 spAspAspAspAspAspTrpProAspSerProProGlnSerAla 91
    : : : : : : : : : : : : : : : : : : : : : : : : :
637 .....CG 636
    : : : : : : : : : : : : : : : : : : : : : : : : :
92 ProGlnAlaArgProThrAlaAlaAlaProArgProProGlyProH1sAr 108
    : : : : : : : : : : : : : : : : : : : : : : : : :
635 CCCGATGCGAGACAAACTGGGTATCGTGTGCTGATCTTAACTCTC 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 gProAlaTrpAlaArgGly.....AlaGlyLeuTrpProp 120
    : : : : : : : : : : : : : : : : : : : : : : : : :
585 CGTTCCTCTCGGGTAGAGGCTTTGGCAGGACAGGCCACAGACGCCCA 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 roThrPro.....ProArgAlaProSerAlaPhe 129
    : : : : : : : : : : : : : : : : : : : : : : : : :
535 AATAGCTGCACACGTAGATGTCATCA..... 506
    : : : : : : : : : : : : : : : : : : : : : : : : :
130 ArgArgAlaSerProSerAlaCysAlaSerProArgSerThrTrpArgAl 146
    : : : : : : : : : : : : : : : : : : : : : : : : :
505 .TGTTCATGTGTAGTGTGTGAAGATATTACTGATGACATCCAGC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 aCysAlaCysAspAla.....ArgAlaG 154
    : : : : : : : : : : : : : : : : : : : : : : : : :
456 GTTGCTTAAAGCTGTACCTCCACGACGACGACATCCACCAACTGGCCA 407
    || : : : : : : : : : : : : : : : : : : : : : : :
154 LysArgGlyArgArgSerProProArgProProArgProProArgProPro 170
    : : : : : : : : : : : : : : : : : : : : : : : : :
406 AGTTGACCTTCCCAAGACAGCATCCATGAGATGACGACGCCAAG... 359
    ||| ||| : : : : : : : : : : : : : : : : : : :
171 ArgProProArgPro...ProArgGlyCysAlaSerArgProThrSerG1 186
    : : : : : : : : : : : : : : : : : : : : : : : : :
358 ....CACATCATGTGGCCAGCGCATACTGAACAGTGTGATGACCACTT 313
    : : : : : : : : : : : : : : : : : : : : : : : : :
186 yCysAlaThrTrpTrp..... 191
    : : : : : : : : : : : : : : : : : : : : : : : : :
312 CCCGAGGAGATCAGCTCAGGAAAGCCCTCCAGCAGAGATTGGCCATGCA 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 .....SerGlyProArgProProAlaTrpArgAlaAla 202
    : : : : : : : : : : : : : : : : : : : : : : : : :
262 CACCAAGGCCAGCATGAGAGGTTGAAGGCCACACTGCTCCAGCTGTGT 213
    : : : : : : : : : : : : : : : : : : : : : : : : :
203 AlaArgGlyProAla..... 207
    : : : : : : : : : : : : : : : : : : : : : : : : :
212 CTCGCGAAGATCGAGGTGAGGAGGCCAAGCCAAATGGCCCATCAGGT 163
    ||| ||| : : : : : : : : : : : : : : : : : : :
208 .SerGlyProThrGlyLeuGlySerGlyAlaGlyTrpArgArgProArgA 224
    : : : : : : : : : : : : : : : : : : : : : : : : :
162 CAGATCTTGGCCAACTGTAGATGATGCCAGAGCC..... 128
    : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

224 rgSerSerGlyArgAla...TrpGlyProArgProValProGlyProTrp 239
    : : : : : : : : : : : : : : : : : : : : : : : : :
127 CCTTTATCCTTAAGGAGAGCGATAGT 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 ProAlaGlnProAlaArgArgTrpArgSer 249
    : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:PCT-US91-06532-3

seq_documentation_block:

```

; Sequence 3, Application PC/TUS9106532
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
; TITLE OF INVENTION: Vaccines and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza Suite 2100
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06532
; FILING DATE: 19910910
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27373/8235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-06532-3

```

alignment_scores:

Quality	124.00	Length:	312
Ratio:	1.033	Gaps:	19
Percent Similarity:	38.462	Percent Identity:	24.679

alignment_block:

US-09-600-714-41/rev x PCT-US91-06532-3 ..

Align seg 1/1 to: PCT-US91-06532-3 from: 1 to: 355

```

921 CCCGAGGAGATCAGCCGACAGCAGACCCAGCACCATTGCGACGCCAG 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 ProthrglyAlaValProthrlaGlInserGlnValThrSerThrProAs 34
    : : : : : : : : : : : : : : : : : : : : : : : : :
871 GAGAGGATCAGGTGACACGAGGTACCCAGCAGCAGCAGCCTCTGCAAC 822
    : : : : : : : : : : : : : : : : : : : : : : : : :
34 nserGlnProAlaVal.ArgSerAlaProAlaAlaAlaProPro.ProPr 50
    : : : : : : : : : : : : : : : : : : : : : : : : :
821 ACCGCACTGTGCACATTAAGTCTTGCTGATCTTCCTTGGGGGTAGCCA. 773
    ||| ||| : : : : : : : : : : : : : : : : : : :
50 oPro.....AlaGlyGlyProProp 57
    : : : : : : : : : : : : : : : : : : : : : : : : :
772 .....AGATGACCCCTGAGA...TGGCTGTACACCGCTGACTGCTA 734
    : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

57 roserCysSerLeuLeuArgIntPLeu.Hls..... 68
733 CAGCATAGTAGTGTGAACAGCGCATTCCTTCATTCAGTTCGCTC 684
69 .....ValPro...GluSerAlaSerA 75
683 AGCAGACGAGATTGAACCTGGCCAGAACATCCACAGAGAGAGG... 638
75 spAspAspAspAspAspAspTrpProAspSerProProGluSerAla 91
637 .....CG 636
92 ProGluAlaArgProThrAlaAlaAlaProArgProProGluProAlaArg 108
635 CCCAGCATGCGAGAACAACTGGTATGCTGCTGCTGATTCATTCCTC 586
108 gProAlaTrpAlaArgGly.....AlaGlyLeuThrProP 120
585 CGTTCCTCGGCTAGAGGCTTTGGCAGGACAGCCACAGAGCCCAA 536
120 rothPro.....ProArgAlaProSerAlaPhe 129
535 AATAGGCTCGACACAGTAGATGTCATCA..... 506
130 ArgArgAlaSerProSerAlaCysAlaSerProArgSerThrTrpArgAl 146
505 .TGTTTCATGCTGTGCTGTTGAAGTATTACTGATGACCATTCCTGAG 457
146 aCysAlaCysAspAla.....ArgAlaG 154
456 GTTCGCTTAACCTGTACCTCCACCAGCAGCATGACACCACTGCCCA 407
154 LysArgGlyArgArgSerProArgProArgProArgProArgProPro 170
406 AGTTGACCTTCCCAAGACAGATCCACTGAGATGAGCAGCAGAACG. 359
171 ArgProProArgPro...ProArgGlyCysAlaSerArgProThrSerG 186
358 ....CACTCATGTGGGCCAGCGCAATACTGAACAAGTGTGATGACCACTT 313
186 yCysAlaThrTrpTrp..... 191
312 CCCAGAGGAGAACTGCTCAGAGAACCGTCCAGAGATTGCCCATGCA 263
192 .....SerGlyProArgProProAlaTrpArgAlaAla 202
262 CACCAAGCGCCAGCATGGAAGAGTTGAAGGCCACACTGCTCCAGCTGTGT 213
203 AlaArgGlyProAla..... 207
212 CTCGGAAGAACTCGAGGTGAGAGAGCCCAATGCGCCCATCATCAGGT 163
208 .SerGlyProThrGlyLeuGlySerGlyAlaGlyTrpArgArgProAla 224
162 CAGATCTTGCCCAACTGTGATGATGCCAGCAGCC..... 128
224 rGserSerGlyArgAla...TrpGlyProArgProAlaProGlyProThr 239
127 CCTTTTGATCTCTAAGAGACCGTCATAGT 98
240 ProAlaGluProAlaArgArgTrpArgSer 249
seq_name: /cgn2_6/protdata/2/laa/5B_COMB.pep:US-08-635-121-2
seq_documentation_block:
; Sequence 2, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-635-121-2

alignment_scores:
Quality: 110.50 Length: 412
Ratio: 0.547 Gaps: 26
Percent Similarity: 49.029 Percent Identity: 21.602

alignment_block:
US-09-600-714-41/rev x US-08-635-121-2 ..
Align seg 1/1 to: US-08-635-121-2 from: 1 to: 1346

1155 ACCTGTCAGAGACGACGAGAGAGAGTATCAGATGCGCCAACTGAGTT 1106
||||| : : : : :
725 ThrcysGlnValIleLysLeuGlnGlnLeuMetGlu..GlnAlaValAl 740
1105 CCCCAGTCTGAGAGAG.....GACTGGAAGCCATCATTCCTCC... 1069
740 aProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrP 757
1068 ..ATTGCCGCTCGCAGCGTATC.....AGCACCAAGCAGACAAATGTA 1027
757 rGluValAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 773
1026 GATGATCTC...TCCAGCAGACCC.....CAGCAAGCTGAAGTTGT 989
774 AspSerGlnAspSerLysAlaThrAlaAlaValAlaValArgGlnSerGlnValTh 790
988 AGCCATGATGAGCTGTGGGAAATCCCGACGACTCGGTTACAACACCCC 939
790 rGluGluGlnAlaAlaThrAlaGlnGlnGluGluProSerThrLeuProA 807
938 GGCAGGACTTGCTGCCCGCCGAG...GGAGATCAGCCCGAGC..... 901
807 snAsnValProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 823

```



```

377 .....TpglyProArgAspSer 382
165 CGTGTGGCGCCATTGGCTTGGCTT..... 191
383 MetAspIyglu.....LeuGlyLeuAspSerIyProSerAspSerle 397
192 .....CCTCACC 198
397 uLeuMetValIyAsnProProProAlaProProGlnProGlnProGlnAla 414
199 TCGAGTTCCGAGACACAGCTGGAGCAGTGGCTTCAACCTTTCAT 248
414 rGlnIyProProGlnProGlnLeu...GlnSerGlnProGlnProGlnSer 429
249 GCTGGCGCTTGTGGAGTGGCAATCGCTGGAGCGCTTCTCTGAGCC 298
430 LeuProProIleAlaValAlaGlnAsnProGlyIy.....ProProse 444
299 AGTTCCTTCTGGAGAGTGGTCACTCACTGTTCAGTATTTCGGCTGGCC 348
444 rArgGlyLeuLeuAspAspIyPheGlnValLeuProGlyIyGlnArgIyS 461
349 ACCATGAGTGTTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
461 erGlyGlnAlaProProGlyIyAspArgSerThrGlyIyGlyIyGly 477
387 TGTCTGGGAGAGTCAACTGGCGCAGTGGTGTGATGGTGGTGGTGG 436
478 AspIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIy 494
437 AGGT..... 440
494 yGlyCyGlySerCySyrProGlyIyLeuArgArgSerLeuLeuLeuH 511
441 .....GACAGCTTGGAGCACTGAGCACTGCTGCTGCTGCTGCTGCTG 476
511 IsGlyAlaArgSerIySyrProIySerCySyrProGlyIyGlyIySerPhe 527
477 TATCTCAACACAGATACACATGATGATGATGATGATGATGATGATG 526
528 GLyAlaArgIySerLeuIleIleHisIySerHisIySerHisIySer 544
527 CAGCCTATTGGGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 576
544 gProIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIy 551
577 GAGGAGACGAGGATAAAGATCAGACAGCAACGATACCCAGTTGTCTGC 626
552 .....GluIySerPheAsnCyS 557
627 CATGCTGGGCGCCCTTCTTGTGTGATGATGATGATGATGATGATGATG 676
558 HisSerIy.....LeuIleArgHisGlnMetH 567
677 C...TCTGTGAGAGTCCATCGAAGAGAAATGC..... 710
567 rHisArgIyGlyIyArgProIyIyIyCySyrSerGlyIyGlyIySerThrIyS 584
711 ..CGTGTCAACACCTACTATGCTGTAGCAGTCAAGCTGTGGAGACCAT 758
584 erArgIySglnHisLeu.....GlnAsnHis 592
759 CTCAGGGATCATCTGGCTCACCCCAAGGAGATCAGCAAGACTATG 808
593 GlnArgLeuHisThrGlyIyGlnArgProPheGln.....Cy 604
809 TGCACAGTGGCGT 821
604 sAlaLeuCySgIy 608
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-933-803A-16
seq_documentation_block:

```

```

; Sequence 16, Application US/08933803A
; Patent No. 6218522
; GENERAL INFORMATION:
; APPLICANT: Saiga, Akihiko
; APPLICANT: Orita, Satoshi
; APPLICANT: Igarashi, Hisanaga
; APPLICANT: Okumura, Kouichi
; APPLICANT: Sakaguchi, Gaku
; TITLE OF INVENTION: DNA MOLECULE RELATING TO SUPPRESSION OF
; GENE EXPRESSION AND NOVEL PROTEIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,803A
; FILING DATE: 19-SEP-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: SHGN-12CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 671 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-933-803A-16

alignment_scores:
    Quality: 108.50      Length: 321
    Ratio: 0.912         Gaps: 16
    Percent Similarity: 37.072    Percent Identity: 23.053

alignment_block:
    US-09-600-714-41 x US-08-933-803A-16 ..

Align seg 1/1 to: US-08-933-803A-16 from: 1 to: 671

42 CCTCTGGCGCCCTAAGCTGAGAC.....AGCTTCATTCTCCTTCT 85
   ||| |||||: ||| |||: |||: |||
350 ProThrAspProAsnSerGlnSerLeuIleSerAlaHisAspIleLeuSe 366
86 ATTTT.....TACCACATATGACGCTTCTCTAGAG 117
   : : : : : |||||
366 rTPIleIySglnIyGlnIyGlnProIyPro..... 376
118 GATCAAAAGGGGCTGTGAGATCCTATCAAGTTGG...CCAGATCTGAC 164
   ||| |||||: ||| |||: |||: |||
377 .....TpglyProArgAspSer 382
165 CGTGTGGCGCCATTGGCTTGGCTT..... 191
383 MetAspIyglu.....LeuGlyLeuAspSerIyProSerAspSerle 397
192 .....CCTCACC 198
397 uLeuMetValIyAsnProProProAlaProProGlnProGlnProGlnAla 414

```



```
296 rSerLeuLeuGlnValThrLeuGlnValLeuIleProArgLeuAlaValP 313
452 GCAACGTGAGATGTCATCATGATATCTTCACACAGACTACCATG 501
313 heSerIleAsnGlnValLeuAlaGlnValVal.....LeuPheMet 326
502 AACATGATGCACATCTACGTGTCGACGCCATTGTTGGG..... 540
327 CysMetAsnThrAlaGlyIlePheIleSerTyrLeuSerAspArgAlaG 343
541 .....CTGTCTGTGGCTTGGCTGCTGCCAAAGCCTTACCCG 577
343 nArgGlnAlaPheLeuGlnThrArgArgCysValGlnAlaArgLeuArg 360
578 AGGACAGGAGATAAAGATCAGACAGACAGATACCCAGTTGTGTGCC 627
360 eugIuThrGlnAsnGlnArgGlnGlnIuArgLeuValLeuSerVal..... 374
628 ATGCTGGGCGCCCTCTTGTGGATGTTTGGCCAAATTTCACACTGTC 677
375 .....LeuProArgPhe..... 378
678 TCTGTAGAAAGTCCAAATCGAAGAGAAATGCCGTGTTCAACACTACT 727
378 ..... 378
728 ATGCTGACAGTACGGGTGTGACAGCCATCTACGGGTCTGCTGGCT 777
379 .....ValValLeuGlnMetIleAsnAspMetThrAsnValGlnAspGlu 393
778 CACCCCAAGGAGATCAGACAGCACTTATGTGCACAGT..... 816
394 HisLeuGlnHisGlnPheHisArgIleTyrIleHisArgTyrGlnAsnVa 410
817 .GCGGTGTGGCAGAGAGCGTGGCTGTGGGTAACCTCGTACCTGATCC 865
410 IserIleLeuPheAlaAspValIysGlyPheThrAsn..... 422
866 CTTCGCGTGGCTTGCATGGTGTGGGCTTGTGGCGCTGATGCTCC 915
423 .....LeuSerThr 425
916 GTCCGGGAGCCAACTACCTGCCGGGTGTGTAAACCGAGTCTGGGG.. 963
426 ThrLeuSerAlaGlnGlnIuLeuValArgMetLeuAsnGlnIuPheAlaAr 442
964 .....ATCCCAACAGCTCATCATGGCTACAACTTCAGCTTGC 1003
442 gPheAspArgLeuAlaHisGln..HisHisCysLeuArgIleLysIleLe 458
1004 TGGGTCTGCTTGG.....AGAGATC 1023
458 uGlyAspArgTyrTyrCysValSerGlyLeuProGluProArgGlnAspH 475
1024 ATTCATATTGTCTGCT.....GTTGCTTGA 1049
475 IAlaHisCysCysValGlnMetGlyLeuSerMetIleLysThrIleArg 491
1050 TACCGTCGAGCGCCGCAATGCGCATGAT 1076
492 pheValArgSerArgThrLysHisAsp 500
```

seq_name: /cgn2_6/prodata/2/laa/5b_COMB.pep:us-08-149-097D-35

seq_documentation_block:

; Sequence 35, Application US/08149097D

; Patent No. 5874236

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

```
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESS: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-55038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0099
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2509 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
FEATURE:
OTHER INFORMATION: /product="Alpha1a-1 subunit of
human calcium channel"
```

US-08-149-097D-35

alignment_scores:
 Quality: 106.50 Length: 382
 Ratio: 0.657 Gaps: 20
 Percent Similarity: 42.408 Percent Identity: 23.560

alignment_block:

US-09-600-714/rev x US-08-149-097D-35 ..

Align seg 1/1 to: US-08-149-097D-35 from: 1 to: 2509

```

1023 GATCTCTCCAGACAGCCAGACAGCTGATGTTAGCC.....CANGA 980
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2185 AspleuSerMetThrThrlSerGlyAspleuProSerGlySglArgAs 2201
979 TGGAGCTGTGGGAAATCCCGAG..... 958
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2201 pGlnGluArgGlyArgProGlyAspArgGlyHisArgGlnHisHisHis 2218
957 .....CAGTCGGTTACACACCCCGGAGAGTACTGCTCCGCC 919
|||  |||  |||  |||  |||  |||  |||  |||
2218 ISHISHISHISHISHISHISHISHISHISHISHISHISHISHISHISH 2228
918 GAGCGAGATGACCCAGCCAGACAGCCAGATGCGACAGCCAGAG 869
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2229 AspLysAspArgGlyArgGlnGluArgProAspHisGlyArgGalaArg 2245
868 AAGGATCAGGTGACAGAGAGTACCAGCCAGCCAGCTCCGCAACAC 819
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2245 ArgAspGln.....ArgTrpSerArgSerProSerGlnGlyArgGlu 2259
818 GCAGTCGTGCACATAGCTCTGCTGATCTCCCTGGGGGTGAG..... 776
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2260 HisMetAlaHisArgGlnGlySerSerSerValSerGlyProAlaP 2276
775 .....CCAAGATGACCTTGAGATGCTG... 752
7276 roSerThrSerGlyThrSerThrProArgArgGlyArgArgGlnLeuPro 2292
751 .....TCACACGCTGACTGCGACAGACATGAGTAGTGTGGAAC 714
2293 GlnThrProSerThrProArgProHisValSerGlySerProValIleAr 2309
713 ACGGATCTCTTCCTTCATGATGAGACTTCTCAGACAGACAGATGGAAC 664
:::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
2309 gLysAlaGlySerGlyProProGlnGlnGlnGlnGlnGlnGlnGlnG 2326
663 TGGCCAGAACATCCACAGAGAGAGGCGCCAGCATGGCAGCAAACTGG 614
:::  ::::  |||||  |||||  |||||  |||||  |||||  |||||
2326 InGlnGlnAlaValAlaArgProGlyArgGalaAlaThrSer..... 2339
613 GATCGGTGCTGTGATCTTATTCCTCCGTCCTCGGTGAGAGGCTTT 564
2340 .....GlyProArgArgGlyTrpProGlyProThrAlaGluProLe 2352
563 GGCAGGACAGAGCCAGACAGACAGCC..... 539
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2352 uAlaGlyAspArgProProThrGlyGlyHisSerSerGlyArgSerProA 2369
538 .....CAAAATAGGCTG.....CGAAGCAGTAGATGTGCATCA 506
:::  ::::  |||||  |||||  |||||  |||||  |||||  |||||
2369 rgMetGluArgArgValProGlyProAlaArgSerGlySerProArgAla 2385
505 TGTTCATGTGTGATGCTGTGTGAGATATTTACTATGACCATCTCTCAG 456
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2386 CysArgHisGlyGlyAla.....ArgTrp..... 2393
455 TTGCCTAAAGTGTGACCTCCAGACAGATGACACACCACTGGGCCAA 406
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2394 .....ProAlaSerGlyProHisValSerGlnG 2403
405 GTTGACCTTCCCAAGACAGATGAGATTCAGACCGCAAAAGCAC 356
::  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

2403 LypProGlyProArgHisHisGlyTyr..... 2412
355 TCATGTGGCCAGCCGAATACGAAACAGTGTGATACACCTTCCAGAA 306
:::  ::::  |||||  |||||  |||||  |||||  |||||  |||||
2413 ...TyrArgGlySerAspTyr.....AspGluAlaAs 2422
305 GGAACCTGGCTCAGAGAGCCGTCACAGAGATTTGCCACTGCACACCAAG 256
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2422 pGlyProGlySerGlyGlyGlnGlnAlaMetAlaGlyAlaTyrAspA 2439
255 CCCAGACATGAAGAGCTTGAAGCCCA.....CACTGCTCCAGCTGTCTC 212
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2439 laProPro.....ProValArgHisAlaSerSerGlyAla 2450
211 TCCGGAACCTGAGGTGAGAGAGCCCAAGCCCAATGGCCGACATCAGGTC 162
:::  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2451 ThrGlyArgSer..... 2454
161 AGATCTTGCCCAACTTGATAGATGACACAGACC.....CCTTTGATC 118
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2455 .....ProArgThrProArgAlaSerGlyProAlaCysA 2466
117 CTCTAAGAGAGCTGATAGTGGTTAAAAAATGAGAGAGAGATGACAG 68
:::  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2466 laSerProSerArgHis.....GlyArgArgLeuProAsnGly 2478
67 CTGCTTCGAGTGTAGG.....CCACAGAGGGGAGGAGCAGC 32
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2479 TyrTyrProAlaHisGlyLeuAlaArgProArgGlyProGlySer 2493

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB pep: US-08-313-288B-15
seq_documentation_block:
: Sequence 15, Application US/08313288B
: Patent No. 5750502
: GENERAL INFORMATION:
: APPLICANT: Jessell, Thomas M. and Avihu Klar
: TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
: TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,288B
: FILING DATE: January 5, 1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0526
: TELEX:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 469 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHEICAL: NO
: ANTI-SENSE: NO

```

US-08-313-288B-15

alignment_scores:

Quality: 105.50 Length: 468
 Ratio: 0.651 Gaps: 27
 Percent Similarity: 34.615 Percent Identity: 21.368

alignment_block:

US-09-600-714-41 x US-08-313-288B-15 ..

Align seg 1/1 to: US-08-313-288B-15 from: 1 to: 469

```

51 CCTAACAGTGAAGCAGCTCTCATCTCTCTTATTTTATTTTACCACCT 100
   ||| ||||| ||| ||| ||||| ||| ||||| ||| ||||| |||
23 ProAlaThrGlySer.....AspProValLeu.CysPheThrGlnT 36
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 ATGACGCTTCTTAGAGATCAAAAGGCGCTGTCGATCCTATCAAGTT 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 YrGluGluSerSerGlyLysCysGlyLeuLeuGlyGlyGlyValSer 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 GGGCAAGATTCAGCGTATGCGGCGCATGGCTTG.....GG 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 ValGluAspCysCysLeuAsnThrAlaPheAlaTyrGlnLysArgSer 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 CTTCCTCAGCTGAGTTCCGAGAGACAGAGCTGAGAGTGGCGCTTCA 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 YrGluLeuGlnProCysArgSerProArgTyrPser..LeuTyrPserThr 85
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 ACCCTCTCATGC..... 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 TrpAlaProCysSerValThrCysSerGluGlySerGlnLeuArgTyrAr 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 .....TGGCGTTGGTGGCGAGTGGG..... 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 GArGcysValGlyTyrPasnGlyGlnCysSerGlyLysValAlaProGlyT 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
272 .....CAATCCCTGAGACGCGCTTCTCTGAGCCAGTTC 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 hrLeuGluTyrProlnLeuGlnAlaCysGluAspGlnGln..CysCysPro 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 CTTCGGGAGAGGTGTCATACACACTTTCAGTATTCGGCTGGCCACCATG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 GluMetGlyGlyTyrPser.....GlyTyrGlyPro.. 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 AGTGTTCGTGGTGTGATGTCAGTGTGATGCTGTCTGGGAAAGTCAA 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 ..... 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 CTGGGCGAGTTGGTGTGATGTCAGTGTGATGTCAGTGTGAGTTCAGCA 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 .....TyrGluProCys..... 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 ACCTGAGATGTCATCAGTATATCTTCACACACAGACTACACATGAAAC 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 .....SerValThrCysSerLysGlyTyrPheTyrThr..... 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 ATGATGACATCTACGTGTGTCAGACCTATTTTGGGCTGTGTGGCTGG 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 ....ArgArgArgAlaCysAsnHisProAlaProLysCys..GlyLysHis 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
555 GTGCTGCCAAAGCCTTCACCCGAGGAGGAGGATAAAGATCAGACAG 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
174 CysProGlyGlnAlaGln.....GluSerGluAla 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
605 CAAGCATACCCAGTTGTCTGC.....CATGC..... 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 acysAspThrThrGlnGlnValCysProThrHisGlyAlaThrProAlaThrTyrPG 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
682 .....TGGGCGCGCTCTCTGTGTGTGATGTCTGGCCAGTTCACACTCG 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 TyrProTyrThrProCysSerAlaSerCysHisGlyGlyProHis..... 214

```

```

677 CTCTCTGAGAAGTTCATCGAAAGAAATGCCGCTGTCAACACCTAC 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 .....GluProLysGluThrArgSerArgLysCysSerAlaPro.. 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
727 TATGCTGTAGACAGTCAGCGTGGTGCACAGCATTCTCAGGGTACCTTGGC 776
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
228 .....GluProSerGlnLysProPro..Gly 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 TCACCC..... 782
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 LysProCysProGlyLeuAlaTyrGlnGlnArgArgCysThrGlyLeuPr 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
783 .....CAAAGGAAAGATCAGCAAGACT 804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
252 oProCysProValAlaGlyTyrGlyTyrGlyProTyrGlyProValSerProC 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
805 TATGTGACACAGTCGCGTGTGGGACAGAGCGCTGTGGTACCCTCGTG 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
269 ysProValThrCysGlyLeuGlnGlnThr..MetGluGlnArgThrCysAs 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
855 TCACCTGATCCCTCTCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 889
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
285 nhisProValProGln..HisGlyGlyProPheCysAlaGlyAspAlaThr 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
890 .....TG 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 ArgThrHisIleCysAsnThrAlaValAlaProCysProValAspGlyLysTr 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
892 GGTCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 916
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 rAspSerThrProLysLysTyrPserProCysIleArgArgAsnMetLysSerI 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
917 .....TGGGAGGAGCAGT.....ACCGCCG 940
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
335 IeSerCysGlnGluIleProGlyGlnGlnSerArgGlyArgThrCysArg 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
941 GGTGT.....GTAC 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 GlyProLysPheAspGlyHisArgCysAlaGlyGlnGlnAlaPheArg 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
952 CGAGTGTGGGAGTATCCACACAGTCCA.....TCATGGGCTACAA 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 gHisCysTyrSerIleGlnHisCysProLeuLysGlySerTyrPserGlyT 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
993 CTTCAGTTCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 rPserThrTyrProLysCysMetProProCysGlyProAsnProThrArg 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1022 TCATCTACATTGTGCTGTGCTGTGCTGTGATACCTGGAGCCG 1063
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 AlaArgGlnArgLeuCysThrProLeuLeuProLysTyrPro 415

```

seq_name: /cgn2_6/prodata/2/1aa/5a_COMB.pep:US-08-466-033-183

seq_documentation_block:

; Sequence 183, Application US/08466033

; Patent No. 5766840

; GENERAL INFORMATION:

; APPLICANT: Kim, Jongsun P.

; APPLICANT: Wages, John

; APPLICANT: Young, Lavonne M.

; APPLICANT: Fry, Kirk E.

; APPLICANT: Linen, Jeffrey M.

; TITLE OF INVENTION: Hepatitis G Virus and Molecular

; TITLE OF INVENTION: Cloning Thereof

; NUMBER OF SEQUENCES: 277

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Ave., Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,033
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/389,886
FILING DATE: 15-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/357,509
FILING DATE: 16-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,729
FILING DATE: 26-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,271
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,558
FILING DATE: 03-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,543
FILING DATE: 03-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,985
FILING DATE: 20-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 2910 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-033-183

alignment_scores:
Quality: 105.00 Length: 507
Ratio: 0.493 Gaps: 29
Percent Similarity: 42.012 Percent Identity: 21.696

alignment_block:
US-09-600-714-41 x US-08-466-033-183 ..
Align seg 1/1 to: US-08-466-033-183 from: 1 to: 2910

10 AAGTACCCGCGTGTCTGCGGCGCTGCGCCCTGCGGCGCTGACT 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1440 LysAlaProAlaGlyValAlaIArgSerGlyProValItrPser..AlaVa 1455
60 GGAAGCA.....GCTCTCA 73
||| |||
1455 IGluAlaGlyValThrTrpTyrGlyMetGluProAspLeuThrAlaAsnL 1472
74 TTCCTCTTATTTTATTTTACCCACTATGACGCTTCTAGAGATCAA 123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1472 euleuArgLeuTyrAspAspCysProTyrThrAlaAlaValAlaAlaasp 1488
124 AAGGGCGCTGACCTATCAAGTGTGCCAAGATCTGACCGTATGAGC 173
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1489 IleGlyGluAlaAlaValAlaPhePheSerGly..... 1498
174 GGCCATTGGCTTGGCTTCACCTCGAGTTTCCGAGACAC..... 216

```

```

1499 .....LeuAlaProLeuArgMetHisProAspV 1508
217 ..ACGTGAGCAGTGTGGCTTCAACCTTCATCCGCTGCTGTGTG 264
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1508 alSerTrpAlaIysValAlaGlyValAsnTrpProLeuLeuValGlyVal 1524
264 ..... 264
1525 GluArgThrMetCysArgGluThrLeuSerProGlyProSerAspAspPr 1541
265 ..CAGTGGCAATCCGCTGAGACGGCTTCTGACGACGATTCCTTC 308
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541 oGlnTrpAla...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGAGAGTGGTCATTCACACTGTTCAGATTCGCTGCGCACATG 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1556 uArgTrp.....GlyAsnAspLeuProSerLysValAlaGlyHisHis 1570
355 AGTGCTTTGTGCGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1571 .....Ile.ValAspAspLeuValArgArgLeuG1 1580
405 CTGGCGCAGTTGGTGTG.....ATGTCCTGTGTG 436
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1580 yValAlaGluGlyTyrValArgCysAspAlaGlyProIleuMetValG 1597
437 AGGTGACAGCTTTAGCAACCGTAGAGATGTCATCGAATATATCTCAAC 486
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1597 lYleuAlaIleAlaGlyMet..... 1604
487 ACAGACTACACATGACATGATGACATCTACGTTCGACGACCTATT 536
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1605 .....IleTyrAlaSerTyrTh 1610
537 TGGCTGTCTGTGGCC.....TGG..... 555
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1610 rGlySerLeuValValThrAspTrpAspValLysGlyGlySerP 1627
556 .....TGCCTGCCAAGACCTCTCCACCCGAGGA 582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1627 rOleuTyrArgHisGlyLysPheGlnAlaThrProGlnProValValGln... 1642
583 ACGAGAGATAAAGATCAGACAGACAGATACCATGTTGTCTGCATGCT 632
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1643 .....ValProProValAspHisArg.. 1649
633 GGGCGCCCTTCTTGTGATGTTCTGGCCAGTTTCACTGCTGCTGC 682
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1650 .....ProGlyGlyGlnSerAla.... 1655
683 TGAGAAATCCATCGAAGAAGATGCCGCTTCAACACCTACTATGCT 732
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668
733 GTACGATC.....AGCTGGTGAACGACCATCTGACGGTCAATC 770
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1669 IleGlnValAspCysAspTrpSerValMetThrLeuSerIleGlyGluVa 1685
771 CTGGCTCACCCCAAGGAGAGATCAGCAAGATTAATGTCACAGTGGC 819
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1685 lleuSerLeuAlaGlnAlaLysThrAlaGlyValAlaLysThrAlaArgAlaL 1702
819 ..... 819
1702 yStrPLeuAlaGlyCysTyrThrGlyThrArgAlaValAlaProThrValSer 1718
820 .....GTGTGGCAGAGAGC...GTGGCTGTGGTACCTGCTG 854
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1719 lIleValAspLysLeuPheAlaGlyGlyTyrTrpAlaAlaValAlaGlyHisC 1735
855 TCACCTGATC.....CCTTCTC 871
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


355 AGTCCTTTGTCGTCGATCTAGTGTGCTCTTTGGGGAAGTCAA 404
1571leValasprleuValargArgleuG 1580
405 CTMGCCAGTGTGTGNG.....ATGCTGCTGGGG 436
1580 yValalagluGlytYrValargCysaspAlaGlyProIleleuMetValG 1597
437 AGTGACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATCTTCAAC 486
1597 LyLeuAlaIleAlaGlyGlyMet..... 1604
487 ACAGACTACACATGACATGATGACATCTACGTGTCCGAGCCTATT 536
1605IleTyralaserTythr 1610
537 TGGGCTGTCTGGCC.....TGG..... 555
1610 rGlyserleuValValThrAspTTPaspValLysGlyLysSerP 1627
556TGCCGCAAGCCTCTACCGAGGGA 582
1627 rOleuTyArgHisGlyAspAlaIleThrProGlnProValValGln... 1642
583 ACGGAGATAAAGATCAGACAGCAACGATCCAGTTTGTCTGCCATGCT 632
1643ValProProValasprHisArg.. 1649
633 GGGGCCCCCTTCTTGTGGATGTTCTGCCCAAGTTTCACTCTGCTGTC 682
1650ProGlyGlyGluSerAla... 1655
683 TGAGAGTCATGGAAGAAGATGCCGTCTCAACACTACTATGCT 732
1656ProserAspAlaLys...ThrValThrAspAlaValAlaAla 1668
733 GTACCACTC.....AGCGTGTGACAGCCATCTAGGTCATC 770
1669 IlegInValasprCysAspTTPserValMetThrleuSerIleGlyGluVal 1685
771 CTGGGCTCACCCCAAGGGAAGATCAGACACTTATGTGCACAGTGGC 819
1685 IleuSerleuAlaGlnAlaLysThrAlaGluAlaTyThrAlaThrAla 1702
819 819
1702 ySTrPleuAlaGlyCysTyThrGlyThrArgAlaValProThrValSer 1718
820GTGTGGCAGAGAGC...GTGGCTGTGGGTACCTCGTG 854
1719 IleValasprLysleuPheAlaGlyGlyThrAlaAlaValAlaGlnHisCy 1735
855 TCACCTGATC.....CCTTCTC 871
1735 shIservAlIleAlaAlaAlaValAlaAlaTyGlyAlaSerArgserP 1752
872 CGTGGCTTGCATGGTGGCTGTGTGGCTGTGGCTGTGGCTGTGGG 921
1752 rOProleuAlaAlaAlaAlaAlaSerTytleuMetGlyLeu...GlyValGly 1767
922 GCA.....GCCAAGTACCTGCCGGGCTGTGTAA 950
1768 GlyAsnAlaGlnThrArgleuAlaSerAlaIleuLeuGlyAlaAlaG 1784
951 CCGAGTCTGGGGGATCCCCACAGCTCATATGGGCTACAACTTC... 996
1784 yThrAlaLeuIleTyThrPro.....ValValGlyLeuIleThrMetAlaG 1798
997 ..AGTTCTGCTGGTCTGTGAGAGATCATCTACATGTGTGTGTGGTG 1044
1798 lyAlaPheMetGlyGlyAlaSerValSerProSerleuValThrIleLeu 1814

1045 CTTGATACCGTCGAGCCGGCAATGCATGATGTGGCTTCCAGGTCTC.. 1092
1815 leuGlyAlaValaGlyGlyTyTrpGluGlyValValaAsnAlaAlaSerleuVal 1831
1093CTCAGCATTTGGGAACTCAGCTTGGCCA 1120
1831 lPheAspPheMetAlaGlyLysleuSerSerGluAspLeuTrpTyThrAla 1848
1121 TCGATAGTCTCTCAGCTGTGCTCTC...CTGACAGTTTGTCTCTAAAT 1167
1848 leProValleuThrSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168 CTAAATATGTGAAGCA 1185
1865 leuValleuTyTrserAla 1870
seq.name: /cgn2_6/prodata/2/iaa/5B_COMB.pep:US-08-464-134-183
seq_documentation_block:
; Sequence 183, Application US/08464134
; Patent No. 5849532
; GENERAL INFORMATION:
; APPLICANT: Kim, Jungsuh P.
; APPLICANT: Wages, John
; APPLICANT: Young, Lavonne M.
; APPLICANT: Fry, Kirk E.
; APPLICANT: Linen, Jeffrey M.
; TITLE OF INVENTION: Hepatitis G Virus and Molecular
; TITLE OF INVENTION: Cloning Thereof
; NUMBER OF SEQUENCES: 277
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,134
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,886
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,509
; FILING DATE: 16-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,729
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,271
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,558
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,543
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,985
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880


```

seq_documentation_block:
: Sequence 183, Application US/08461361
: Patent No. 5856134
: GENERAL INFORMATION:
: APPLICANT: Kim, Jungshuh P.
: APPLICANT: Wages, John
: APPLICANT: Young, Lavonne M.
: APPLICANT: Fry, Kirk E.
: APPLICANT: Linmen, Jeffrey M.
: TITLE OF INVENTION: Hepatitis G Virus and Molecular
: TITLE OF INVENTION: Cloning Thereof
: NUMBER OF SEQUENCES: 277
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave., Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,361
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/389,886
: FILING DATE: 15-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/357,509
: FILING DATE: 16-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/329,729
: FILING DATE: 26-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,271
: FILING DATE: 23-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,543
: FILING DATE: 03-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,985
: FILING DATE: 20-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Fabian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880
: TELEFAX: (415) 324-0960
: INFORMATION FOR SEQ ID NO: 183:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2910 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-461-361-183

```

```

alignment_scores:
Quality: 105.00 Length: 507
Ratio: 0.493 Gaps: 29
Percent Similarity: 42.012 Percent Identity: 21.696

```

```

alignment_block:
US-09-600-714-41 x US-08-461-361-183

```

```

Align seg 1/1 to: US-08-461-361-183 from: 1 to: 2910
10 AAGTACCCGCGGTCTGTCGGCGCTGCCCTGCGCCCTGAGCCCTAACACT 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1440 LysAlaProAlaGlyValValArgSerGlyProValTyrPse...AlaVal 1455
60 GGAGCA.....GCTGCA 73
1455 LgluAlaGlyValThrTyrGlyMetGluProAspLeuThrAlaAsnL 1472
74 TTCCTCTCTCTATTTTTCACCACTATGACGCTTCCTCTAGAGATCAA 123
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1472 euLeuArgLeuTyrAspAspCysProTyrThrAlaAlaValAlaAsp 1488
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1489 lIeGlyGluAlaAlaValAlaPhePheSerGly..... 1498
174 GGCAATGGCTTGGGCTTCACCTGACGAGTTCCGAGACAC..... 216
1499 .....LeuAlaProLeuArgMetHisProAspV 1508
217 ..AGCTGAGCAGTGTGGCTTCAACCTTCATGCTGCGCTGTGTG 264
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1508 alSerTyrAlaLysValArgGlyValAsnTyrProLeuLeuValGlyVal 1524
264 ..... 264
1525 GluArgThrMetCysArgGluThrLeuSerProGlyProSerAspAspPr 1541
265 ..CAGTGGCAATCGTGGAGCGCTTCCTGAGCAGTTCCCTTC..... 308
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541 cGlnTyrPala...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGAGAGTGCATCAACACGTTCACATTCAGTTCGGCTGCCACCATG 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1556 uArgTyr.....GlyAsnAspLeuProSerLysValAlaGlyHisHisL 1570
355 AGTCGTTTGTGCTGCTCATCTCAGTGCATGCTGCTTGGGAGAGTCAA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1571 .....Ile.ValAspAspLeuValArgArgLeuGl 1580
405 CTGGCGCAGTTGGTGTG.....ATGTGCTGTGTG 436
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1580 yValAlaGluGlyTyrValArgCysAspAlaGlyProIleLeuMetValG 1597
437 AGGTACACGCTTAGCAACCTGAGAGTGCATCAGTAATATCTTCAAC 486
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1597 lYLeuAlaIleAlaGlyMet..... 1604
487 ACAGACTACACATGACATGATGCATCTACGTGTCGACGCTATT 536
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1605 .....IleTyrAlaSerTyrThr 1610
537 TGGGCTGTCTGTGCC.....TGG..... 555
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1610 rGlySerLeuValValValThrAspTyrAspValLysGlyGlySerP 1627
556 .....TGCTGCCAAAGCCTTACCAGAGGA 582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1627 rGlyLeuTyrArgHisGlyAspGlnAlaThrProGlnProValValGln... 1642
583 ACGAGAGATTAAGATCAGACAGCAACGATACCCAGTTGTGTCATGCT 632
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1643 .....ValProValAlaAspHisArg... 1649
633 GGGCGCCCTCTTCTGTGGATTTCTGGCCAAGTTCAACTGCTGTGC 682
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1650 .....ProGlyGlyGlnSerAla.... 1655
683 TGAGAAGTCAATCGAAGAGAGATGCCGTGTCAACACCTACATATGCT 732
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668

```



```

174 GGCATTGGCTGGGCTTCCTACCTCGAGTTCCGAGACAC..... 216
      ::::: |||
1499 .....LeuAlaProLeuArgMethHisProAspV 1508
217 ..AGCTGGACAGTGTGCGCTTCAACCTCTTCAGCTGGCGCTGTGTG 264
      ||||| ::::: ||| ::::: |||
1508 aISerTPAlaLysValArgLysValAsnTPProLeuLeuValGlyVal 1524
264 ..... 264
1555 GlnArgThrMetCysArgGluThrLeuSerProGlyProSerAspAspPr 1541
265 ..CACTGGGCAATCCTCTGAGCGGCTTCAGCCAGTTCCTCC..... 308
      ||||| ||| ||| ::::: |||||
1541 oGlnTPAla...GlyLeuLysGly..ProAsnProValProLeuLeuLe 1556
309 ....TGGGAAGTGTGTATCAGTCTGATTCAGTATTCGGCTGGCACCCTG 354
      ||| ||||| ::::: |||||
1556 uArgTPR.....GlyAsnAspLeuProSerLysValAlaGlyHisHis. 1570
355 AGTCTTTGTGCGTGTGATCTCAGTGTGCTGTGGGGAAGTCA 404
      :: ||||| ::::: |||||
1571 .....Ile.ValAspAspLeuValArgArgLeuG1 1580
405 CTGAGCGCAGTTGGTGTG.....ATGGTGTGTGG 436
      ::|||::: |||
1580 yValAlaGluLysValArgCysAspAlaG1ProIleLeuMetValG 1597
437 AGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATCTTCAC 486
      ::::: |||:::
1597 LysLeuAlaIleAlaGlyLysMet..... 1604
487 ACAGACTACACATGACATGATGACATCTACGTGTGCGAGCCTATT 536
      ::::: |||||
1605 .....IleTyrAlaSerTyrTh 1610
537 TGGGCTGTGTGGC.....TGG..... 555
      ||| |||||
1610 rGlySerLeuValValAlaThrAspTPAspValLysGlyLysSerP 1627
556 .....TGGCTGCCAAGCCTTCACCGCAGGA 582
      ::|||::: |||||
1627 rOleuTyrArgHisGlyAspGlnAlaThrProGlnProValAlaGln... 1642
583 ACGAGATTAAGATACAGACAGCAACGATACCTGTGTGCCATGCT 632
      ::|||::: |||||
1643 .....ValProProValAspHisArg.. 1649
633 GGGGCGCCCTTCTTGATGTTCTGGCAAGTTTCACTGTGCTGTGC 682
      ||||| ::|||
1650 .....ProGlyGlyLysSerAla.... 1655
683 TGAGACATCCATCGAAGAAGATGCCGTTCACACCTACTATGCT 732
      ||| ::||| ::|||
1656 .....ProSerAspAlaLys...ThrValThrAspAlaValAlaAla 1668
733 GTACAGATC.....AGCTGTGACAGCAGCATCTCAGGTCATC 770
      ::||| ||||| |||||
1669 lIleGlnValAspCysAspTPSerValMetThrLeuSerIleGlyGluVa 1685
771 CTGGGCTCACCCCAAGGAAGATCAGACAGACTTATGTGCACAGTGGC 819
      ||||| ||||| ::|||
1685 lLeuSerLeuAlaGlnAlaLysThrAlaGluAlaTyrThrAlaThrAla 1702
819 ..... 819
1702 yStrPLeuAlaGlyCysTyrThrGlyThrArgAlaValProThrValSer 1718
820 .....GTGTGGCAGAGAGC...GTGGCTGTGGTACTCGG 854
      ::|||::: |||||
1719 lIleValAspLysLeuPheAlaGlyLysTyrPAlaAlaValAlaGlyHis 1735
855 TCACCTGATC.....CCTTCTC 871

```

```

1735 sHisSerValIleAlaAlaValAlaAlaLysGlyAlaSerArgSerP 1752
      ||| ::::: |||
872 CGTGCGCTGCATGCTGCTGGCTGTGGCTGGGCTCATCTCCGTCGG 921
      || ||||| ::::: ||| ::|||
1752 rOProLeuAlaAlaAlaAlaSerTyrLeuMetGlyLeu...GlyValGly 1767
922 GGA.....GCCAAGTACTGCCCGGGGTGTGTA 950
      ||| ::||| |||||
1768 GlyAsnAlaGlnThrArgLeuAlaSerAlaLeuLeuGlyAlaAlaG1 1784
951 CCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTACAACTTC.... 996
      ::||| ||| ::|||
1784 yThrAlaLeuGlyThrPro.....ValValGlyLeuThrMetAlaG 1798
997 ..AGCTGCTGGGTCTGCTGTGAGAGATCATCTACATGTGCTGTGTG 1044
      ::||| |||||
1798 lYAlaPheMetGlyGlyAlaSerValSerProSerLeuValThrLeu 1814
1045 CTGTATACGTCGAGCCGCAATGGCATGTGCTTCACAGTCTTC.. 1092
      ||| ::||| |||||
1815 LeuGlyAlaValGlyGlyTyrPGLuglyValAlaAlaAlaSerLeuVa 1831
1093 .....CTCAGCATTTGGGAACTCAGCTTGGCCA 1120
      ||||| |||||
1831 lPheAspPheMetAlaGlyLysLeuSerSerGlyAspLeuTPyrAlaI 1848
1121 TCGTATAGCTCTCAGCTGTGCTC...CTGACAGCTTGTCTCTTAAT 1167
      ||| ::||| |||||
1848 lProValLeuThrSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168 CTAAATATGGAAGCA 1185
      || ::|||
1865 LeuValLeuTyrSerAla 1870
      || ::|||

seq_name: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.PCT-US95-06266-157
seq_documentation_block:
; Sequence 157, Application PC/TUS9506266
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Detection of Viral Antigens Coded
; TITLE OF INVENTION: by Reverse Reading Frames
; NUMBER OF SEQUENCES: 157
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06266
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,985
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,561
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,729
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,271
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:

```


1121 TCGTGAATAGCTCTCAGCTGTGGTCTC...CTGACAGGTTTGCTCCTAAAT 1167
||::: ||| |||::: |||:::
1848 IeProValLeuThrSerProGlyAlaGlyLeuAlaGlyIleAlaLeuGly 1864
1168 CTTAAATATGGAAGCA 1185
|||::: |||
1865 LeuValLeuTyrSerAla 1870

THIS PAGE BLANK (USPTO)


```

351 CATGAGTCTTTGCGTGATCTCAGTGAGTCTGTTGGGAAG 400
|||||
117 mterSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
|||||
401 TCAACTTGGCCAGTTGGTGTGATGGTGTGGTGGAGGTACACCTTA 450
|||||
134 AlAsnLeuAlaGlnLeuValAlaValMetValLeuValGlnValThrAlaLeu 150
|||||
451 GGCAACCTGAGGATGGTCACTAGTAAATATCTTCAACACAGACTACACAT 500
|||||
151 GlnAsnLeuAlaGlnMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
|||||
501 GAACATGATGACATCTACGTGTTCGACGCTATTTGGGCTGTCTGTGG 550
|||||
167 LAsnMetMetHisIleTyrValAlaPheAlaIleTyrPheGlnValLeuSerVal 184
|||||
551 CCTGTGGCTGCCAAGGCTCTACCCGAGGAGACGAGGATTAATATACAG 600
|||||
184 LATPCysLeuProLysProLeuProGlnGlyThrGlnAspLysAspGln 200
|||||
601 ACAGCAAGATACCCAGTTGTCTGCATGCTGGGCGCCCTCTTCTGTGG 650
|||||
201 ThrAlaThrIleProSerLeuSerAlaMetLeuGlnAlaLeuPheLeuTr 217
|||||
651 GATGTTCTGGCCAAAGTTTCAACTGTCTGTCTGTGAAGATCCAAATCGAAA 700
|||||
217 pMetPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleLysA 234
|||||
701 GGAGAAATGGCGGTGTTCACACACTACTGTGTGACAGTACAGCTGGTG 750
|||||
234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerValVal 250
|||||
751 ACAGCCATCTCAGGGTCACTCTTGGCTCACCCCAAGGAGATACAGAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerLys 267
|||||
801 GACTTATGTGCACAGTGGCGGTGTGGCAGAGCGGTGGCTGTGGTACT 850
|||||
267 sThrTyrValHisSerAlaValAlaLeuAlaGlnGlyValAlaValGlyThrS 284
|||||
851 CGTGACACCTGATCCCTTCCGTTGGCTTGGCATGTGGTGGTCTTGTG 900
|||||
284 eGlySerHisLeuIleProSerProThrLeuAlaMetValLeuGlnLeuVal 300
|||||
901 GCTGGCGATCTCCGTCGGGGAGCCAAAGTACTCGCGGGGTGTGTAA 950
|||||
301 AlaGlyLeuIleSerValGlnGlyAlaLysTyrLeuProGlnLysCysAs 317
|||||
951 CCGAGTGTGGGATTTCCCAAGCTCCATCATGTGGCTACAACTTACACT 1000
|||||
317 nArgValLeuGlyIleProHisSerSerIleMetGlnTyrAsnPheSerL 334
|||||
1001 TGTGTGGTGTCTGTGGAGAGATCATCTACATTTGTCTGTGGTCTTAT 1050
|||||
334 euleuGlnLeuLeuGlnGlyIleIleTyrIleValLeuLeuValLeuAsp 350
|||||
1051 ACCGTCGAGCGCGCAATGCGCATGTGGCTTCCAGTCTCCATCAGAT 1100
|||||
351 ThrValGlnAlaGlnLysGlnGlyMetIleGlyPheGlnValLeuLeuSerI 367
|||||
1101 TGGGAACTCAGCTTGGCCATCGTATAGCTCTCACTGTGTGGTCTCTGA 1150
|||||
367 eGlyIleuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
|||||
1151 CAGGTTTCTCTTAATCTTAAATATATGAAAGACCTCATGAGGCTTAA 1200
|||||
384 hArgIleuLeuLeuAsnLeuLysIleTyrAlaAlaProHisGlnAlaLys 400
|||||
1201 TATTTTGAACAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTT 1250
|||||
401 TyrPheAspAspGlnValAlaPheThrLysPheProHisLeuAlaValGlyPh 417
|||||
1251 T 1251

```

```

seq_name: sp_human:09UQ22
417 e 417
seq_documentation_block:
ID 09UQ22 PRELIMINARY; PRT; 417 AA.
AC 09UQ22;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TO;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of Rhdya in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018967; BAA81900.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA: 45211 MW: 764B092CBF95ECC4 CRC64;

```

```

alignment_scores:
Quality: 2117.00 Length: 417
Ratio: 5.077 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.520

```

```

alignment_block:
us-09-600-714-41 x 09UQ22 ..

```

```

Align seg 1/1 to: 09UQ22 from: 1 to: 417

```

```

1 ATGAGCTTAAGTACCCGCGGTCTGTCGCGGCTGCTGCCCCCTGTGGC 50
|||||
1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTrpAl 17
|||||
51 CCTAACACTGGAAGCAGCTCTCATTTCTCTTATTTTTCACCACT 100
|||||
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
|||||
101 ATGAGCGTCTCTTAGAGAGATCAAAAGGGCTCGTGGCATCTTATCAAGTT 150
|||||
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaIleSerTyrGlnVal 50
|||||
151 GGCCAAATCTGACCGTGTATGATGGCGCCATTGGCTTGGCTTCTACCTC 200
|||||
51 GlnGlnAspLeuThrValMetAlaAlaIleGlyLeuGlnPheLeuThrSe 67
|||||
201 CAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACTCTTCACTGC 250
|||||
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
|||||
251 TGGCGCTGTGTGTCAGTGGGCAATCGCTGTGAGCGGCTTCTGAGCCAG 300
|||||
84 euaAlaLeuGlnGlyValGlnTrpAlaIleLeuLeuAspGlnPheLeuSerGln 100
|||||
301 TTCCCTTCTGGGAAGGTGTGTCATCACACTGTTCAGTATCGGTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
|||||
351 CATGAGTCTTGTTCGTCGTCGATCTCAGTGGATGCTGTCTGTGGGAAG 400
|||||
117 mterSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlnLysV 134

```



```

401 TCAACTTGGCGCAGTTGGTGTGATGGTGTGAGTGAGTGACAGCTTTA 450
134 aAsnLeuAlaGlnLeuValValMetValLeuValGluValThralaleu 150
451 GGCACCTGAGAGATGTCATCATGTAATATCTTCAACACAGACTACCAT 500
151 GlysLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATTCACGTTGCTGCGCATGCTGAGGCGCTTCTGTG 550
167 CAsMetMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerValA 184
551 CTTGGTCCCTGCCAAAGCCTCTACCCGAGGAGGAGGAGATGAAGATCAG 600
184 latrPcysLeuProLysProLeuProGlnGlyThrGlnAspLysAspGln 200
601 ACAGCAAGATACCCAGTTGTGTGTCATGCTGGGCGCTTCTGTG 650
201 ThrIatThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTGTGGCGCAGTTTCAACTGCTGCTGCTGAGAAAGTCCATCGAA 700
217 pMePheTrProSerPheAsnSerAlaLeuAlaArgSerProIleGlnA 234
701 GGAAGATGCCGCTGTCAACACCTACTATGCTGTAGAGTACGCGTGGTG 750
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCATCTCAGGGTCATCTTGGCTACCCCGAGGAGGAGATCAGCA 800
251 ThrIatIleSerGlySerSerLeuAlaHisProGlnGlyHisIleSerly 267
801 GACTTATGTGACAGTGGGCTGTGGAGGAGGCGTGGCTGGTACCT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValGlyTrpS 284
851 CGTGTACCTGATCCCTTCTCCGTGCTGCGCATGAGTGTGCTGTGTG 900
284 eRcysHisIleuIleProSerProTrIlePheAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGCGGAGGAGCCAGTACCTCCGCGGCTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGAGTCCCGCAGCTGCATGAGGCTACAACTTCAGCT 1000
317 natrValLeuGlyIleProHisSerIleMetGlyTyrAsnPheSerL 334
1001 TGCTGGGTCTGCTTGAGAGATCATCTACATTTGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyGlyIleIleTyrIleValIleLeuValLeuAsp 350
1051 ACCGTGGAGCCGCAATGCGATGATGGCTTCCAGGCTCCTCAGCAT 1100
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValIleuLeuSerI 367
1101 TGGGGAATCAGTGGGATCGATAGTACGTCAGGCTGCTGCTCCGA 1150
367 eGlyGlyLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTTGCTCTAAATCTTAAATATGGAAGCAGCTCATGAGGCTAAA 1200
384 hrgIyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlyAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGAGATTTCTCATTTGCTGTGGATTT 1250
401 TyrPheAspArgGlnValAlaPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_name: sp_human:Q9NXY8
seq_documentation block:
ID Q9NXY8 PRELIMINARY; PRT; 417 AA.
AC Q9NXY8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 15, Last annotation update)
DE RHESUS BLOOD GROUP, D ANTIGEN.
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner F.F., Flegel W.A.;
RT "Molecular structure of weak D type 18, another RHD allele detected
RT among weak D phenotypes."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ287289; CAB75731.1; JOINED.
DR EMBL; AJ287291; CAB75731.1; JOINED.
DR EMBL; AJ287292; CAB75731.1; JOINED.
DR EMBL; AJ287293; CAB75731.1; JOINED.
DR EMBL; AJ287294; CAB75731.1; JOINED.
DR EMBL; AJ287295; CAB75731.1; JOINED.
DR EMBL; AJ287296; CAB75731.1; JOINED.
DR EMBL; AJ287297; CAB75731.1; JOINED.
DR EMBL; AJ287298; CAB75731.1; JOINED.
DR InterPro: IPR001905;
DR InterPro: IPR002229;
DR Pfam: PF00909; Ammonium transp; 1.
FT PRINTS: PRO0342; RHESUSRHD.
FT VARIANT 1 1 45210 KM; 60FB4CF816F163EF CRC64;
SQ SEQUENCE 417 AA; 45210 MW; 60FB4CF816F163EF CRC64;

alignment_scores:
Quality: 2115.00 Length: 417
Ratio: 5.084 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.760

alignment_block:
US-09-600-714-41 x Q9NXY8
Align seg 1/1 to: Q9NXY8 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCGCTGCCCTCGGGC 50
1 MetSerSerLysTyrProTrpSerValAlaArgCysLeuProLeuTrpAl 17
51 CCTAACATCTGAGACGCTCTCATTTCTCTATTTTTCACCCACT 100
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheTrpHis 34
101 ATGAGCTTCTCTAGAGAGTCAAAAGGGCTCGTGGCATCTCATCAAGTT 150
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAAGTCTGACCGTGTAGTGGCGCATGCTGGCTTGGCTTCACCTC 200
51 GlyGlnAspLeuThrValMetAlaIleAlaIleGlyLeuGlyPheLeuThr 67
201 GAGTTTCCGGAGACACACACTGGAGAGTGTGGCTTCAACTCTTCATGC 250
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTGTGTGTCAGTGGGCAATCTGCTGAGAGGCTTCTCGAGCCAG 300
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCTGGGAAGTGTGTCATCAGACTGTTCAGTATTCGGCTGGCCAC 350

```

```

|||||
101 pheProserGlyValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGCTCTTGGGTCGATGTCAGTCAGTGCCTGCTGGGAGG 400
117 rMeSerIleuSerValIleIleSerValAspAlaValLeuGlyLysV 134
401 TCAACTTGGCCAGTTGGTGGATGGTGTGTGTGAGGTACAGCTTGA 450
134 alaLeuAlaGlnLeuValValIleMetValLeuValGluValThrAlaLeu 150
451 GGCAACTGAGAGATGTCATGATGATATCTTCAACACAGACTACCAT 500
151 GlyAsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTGTCCGACCTATTGGCTGTCTGTGG 550
167 LAsnMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerValA 184
551 CCTGGTCCCTCCAAAGCCTTACCCGAGGAGACGAGATGAAGATCAG 600
184 latrPCysLeuProIysProIeuProGluGlyThrGluAspLysAspGln 200
601 ACAGCAACGATACCCAGTTGTGTGCCATGCTGGGCGCCCTTCTGTGG 650
201 ThrIatThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTCGGCAAGTTCAACTGCTGCTGCTGTGAGAGTCCATCCGAA 700
217 pMeThrIleProSerPheAsnSerAlaLeuLeuArgSerProIleGluA 234
701 GGAAGATGCGCGTGTCAACACACTACTAGTGTAGACAGTACGCTGG 750
234 rGlyAsnAlaValAlaPheSerThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCCATCTCAGGGTCACTCTGGCTACCCCAAGGAGGAAGATCAGCA 800
251 ThrIatIleSerGlySerSerLeuAlaHisProGluGlyLysIleSerLy 267
801 GACTTATGTGCACAGTGGGCTGGTGGCGAGGAGGCTGGCTGGGTACT 850
267 sThrTyrValHisSerAlaValIleAlaGlyValAlaValGlyThrS 284
851 CGTGTACCTGATCCCTTCTCCGTGGCTTGCATGTGCTGGGTCTTGTG 900
284 eGlySHisLeuIleProSerProIleuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGAGGAGCAAGTACTGCCGGGTGTGTAA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGAGATCCCCACAGCTCCATGGGCTACAACTTCACT 1000
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
1001 TGTGGGCTGTGCTGGAGAGATCATTAACATGTGTGCTGTGCTGTGAT 1050
334 euLeuGlyLeuLeuGlyGluIleIleTyrIleValIleLeuValLeuAsp 350
1051 ACCGTGGAGCCGCAATGGCATGATGGCTTCCAGGCTCTCCAGAT 1100
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValIleLeuSerI 367
1101 TGGGAACTCAGCTTGGCCATGTCGATAGCTTCACGCTGGTCTCCGGA 1150
367 eGlyGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTGCTCTTAATCTTAATATGGAAGAAGCACTCATGAGGCTAAA 1200
384 hArgLyLeuLeuAsnLeuLysIleTyrLysAlaProHisGlyAlaLys 400
1201 TATTTTGTGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
|||||

```

```

401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:09023

```

```

seq_documentation_block:
ID 09023 PRELIMINARY; PRT; 417 AA.
AC 09023;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FK;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of RhDva in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018966; BAA81899.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SO SEQUENCE 417 AA; 45131 MW; 6D5F4D6CA0DAA884 CRC64;

```

```

alignment_scores:
Quality: 2113.00 Length: 417
Ratio: 5.079 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.520

```

```

alignment_block:
US-09-600-714-41 x 09023 ..

```

```

Align seg 1/1 to: 09023 from: 1 to: 417

```

```

1 ATGAGCTCTAGTACCCGGGTGTGTCGGGCGTGGCCCTCTGAGC 50
|||||
1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTrpAl 17
51 CCTAACACTGGAGACAGCTCATCTCTCTCTCTATTTTTCACCACT 100
|||||
17 AleuThrLeuGluAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGACGCTTCTTAGAGATCAAAAGGGCGTCGCGCATCTATCAAGTT 150
|||||
34 YrAspAlaSerLeuGluAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAGATCTGACCGTGANAGCGGCATGGCTTGGGCTTCTCACCTC 200
51 GlyLinspleuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGAGACACACAGCTGAGCAGCTTGCCCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgTrgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGTGTGTCAGTGGGCAATCTGCTGAGACGCGCTTCTTAGAGCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTCTGGAGAGTGTGCATCACACTGTTCAGTATTCGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117

```

```

351 CATGAGTGGCTTTGTCGGTCTGATCTCATGAGTGTCTCTTGGGGAAG 400
|||||
117 rmetserAlaLeuSerValLeuIIeSerValAspAlaValLeuGlyLysV 134
401 TCACCTTGGCGCACTTGGTGTGATGCTGCTGGTGGAGGTACAGCTTGA 450
|||||
134 alaLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCACACCTGGAGATGTCATCATGTAATATCTTCAACACAGACTCCACAT 500
|||||
151 GlysLeuArgMetValIIeSerAsnIIePheAsnIIePheAsnIIePheSwe 167
501 GAACATGATGCACATCTACGTGTTCGACCCCTAATTTGGGCTGTCTGTGG 550
|||||
167 tAsnMetMetHisIleTyrValIlePheAlaIleTyrPheGlyLeuSerVal 184
551 CCTGTGCTCTCCAAAGCTCTTACCCGAGGAGGAGGAGGATTAAGATCAG 600
|||||
184 latrPcysLeuProIysProIysProGluGlyThrGluAspLysPcIn 200
601 ACAGCAACGATACCAAGTTGCTGTCATGCTGGGCGCCCTCTTCTGTG 650
|||||
201 ThrAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGTCTGTCTGAGAAATCCAAATCGAA 700
|||||
217 pMetPheTrpProSerValAsnSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGATGCGCGTGTTCACACACTACTATGCTGTAGCAGTACGCTGTG 750
|||||
234 rGlyAsnAlaValIlePheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCAATCTCAGGGTCTCTTGGCTCACCACCCCAAGGAGGAGATACGAA 800
|||||
251 ThrAlaIIeSerGlySerSerLeuAlaHisProGlnGlyLysIIeSerLys 267
801 GACTTATGTGCACAGTGGGCTGTGGCAGAGGCGCTGGCTGTGGTACT 850
|||||
267 sThrTyrValHisSerAlaValLeuAlaGlyLysAlaValAlaGlyLys 284
851 CGTGTACCTGATCCCTTCTCGTGGCTTGCATGTGCTGTGGTGTGTG 900
|||||
284 erCysHisLeuIIeProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCGATCTCCGTCGGGGGAGCACAAGTACCTGCCGGGGTGTGAA 950
|||||
301 AlaGlyLeuIIeSerValGlyValAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATTTCCACAGCTCCATCATGGGCTACAACTTCACACT 1000
|||||
317 nArgValLeuGlyIleProHisSerIleMetGlyTyrAsnPheSerI 334
1001 TGTGGTGTCTGCTTGGAGATCATCTACATTTGCTGTGCTGTGCTTGA 1050
|||||
334 euleuGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTGGAGCGCGCAATGCGATGATGGCTTCCAGGTCCTCCATGAGAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGAACTCAGCTTGGCCATCGTATGAGCTCTCAGCTGTGCTGCCGTA 1150
|||||
367 eGlyIleuLeuSerLeuAlaIleValIIeAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTCTCTCTAAATCTTAAATATGGAAGACACTCATGAGGCTTAAA 1200
|||||
384 hngIyLeuLeuLeuAsnLeuLysIIeTrpLysAlaProHisGlnAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGGTGGAT 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 r 1251

```

```

seq_name: sp_human:Q9UQ21
417 e 417
seq_documentation_block:
ID Q9UQ21 PRELIMINARY; PRT; 417 AA.
AC Q9UQ21;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE RH BLOOD GROUP D ANTIGEN (RHD).
GN RHDVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBL_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=TT;
RA Hyodo H., Ishikawa Y., Kashiwase K., Ogawa A., Watanabe Y.,
RA Tsuneyama H., Toyoda C., Uchikawa M., Akaza T., Fujii T.;
RT "Polymorphisms of Rhdva in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018968; BAA81901.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45177 MW; B7327B16747E61D CRC64;

alignment_scores:
Quality: 2107.00 Length: 417
Ratio: 5.065 Gaps: 0
Percent Similarity: 99.760 Percent Identity: 99.041

alignment_block:
US-09-600-714-41 x Q9UQ21 ...

Align seg 1/1 : to: Q9UQ21 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCCCTCTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValArgLysLeuProLeuTrpAl 17
51 CCTAACACTGGAGACGCTTCATTTCTCTCTTATTTTTCACCACT 100
|||||
17 alaThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHis 34
101 ATGAGCTCTTCTAGAGATCAAAAGGGCTGTGGCATCCATCAAGTT 150
|||||
34 yTrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAAGTCTGACCGTATGGCGGCATTTGGCTTGGCTTCCACCTC 200
|||||
51 GlysAlaSerPheThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGAGACACACAGCTGGAGCAGTGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgArgHisSerTrpSerSerValAlaIlePheAsnLeuPheMet 84
251 TGGCGCTGTGTGTCAGTGGGCAATCGTGGGAGGCTTCCGAGCGCAG 300
|||||
84 euaIleLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTGGGGAAGGTGTCATCACACTGTTCATATTCGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTGTCTTGTGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAG 400
|||||
117 rMetSerAlaLeuSerValLeuIIeSerValAspAlaValLeuGlyLysV 134

```

401 TCAACTGGCCAGTTGGTGTGATGTCGTGTGAGAGTGACAGCTTTA 450
 134 aAAsnLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
 451 GGCAACCTGAGAGATGATCATGATATATCTTCAACACAGCATACCAT 500
 151 GlyAsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisLe 167
 501 GAACATGATGACATCTAGCTGTCCGAGCCTATTGGCTGTCTGG 550
 167 CasMetCHisIleTyrValPheAlaIleTyrPheGlyLeuSerValA 184
 551 CCTGGTCCCTCCAAAGCCTTACCCGAGGAGACGAGATAAAGATAG 600
 184 IaTrpCysLeuProIysProIeuProGlnGlyThrGlnAspLysAspIn 200
 601 ACAGCAGATACCCAGTTGTCTGCGCATGCTGGCGCCCTCTTGTG 650
 201 ThrAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
 651 GATGTTGGCCAGTTTCAACTGCTGCTGCTGAGAGTCCATCGAA 700
 217 PhePheTrpProSerValAsnSerAlaLeuLeuArgSerProIleGlnA 234
 701 GGAGAAATGCCGTGTTCACACCTACTATGCTGTAGACATGAGCGTGTG 750
 224 rGlyAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
 751 ACAGCATCTCAGGGTCATCTGGCTACCCCAAGGAAAGATCAGCA 800
 251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerL 267
 801 GACTATGTGCACAGTGGGTGTGGAGAGGCGGTGGCTGGGTACT 850
 267 sThrTyrValHisSerAlaValLeuAlaGlyGlyAlaValAlaGlyHis 284
 851 CGTGCACCTGATCCCTTCCGTCGCTGTCATGTCGTCGTCGTCGTCG 900
 284 eRcYHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
 901 GCTGGGCTGATCTCCGTGGGGGAGCCAGTACTGCCGGGCTGTGTA 950
 301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
 951 CCGAGTGTGGGATCCCCACAGCTCATGGGCTACAACTTCACT 1000
 317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
 1001 TGCTGGTCTGCTGGAGAGATCATCAATGTGCTGCTGCTGCTGAT 1050
 334 euLeuGlyLeuLeuGlyGlnIleIleTyrIleValLeuLeuValLeuAsp 350
 1051 ACCGTGCGAGCCGCAATGCGATGATTGGCTCCAGGTCCTCCAGAT 1100
 351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerL 367
 1101 TGGGAACTAGCTGGCCATGATAGTGTGACGTCGTCGTCCTGCA 1150
 367 ecGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
 1151 CAGGTTGGCTCTAAATCTTAAATATGGAAGAAGCCTCATGAGGCTTAA 1200
 384 nArgLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTATTTGGCTGTGGATT 1250
 401 TyrPheAspArgGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
 1251 T 1251
 417 e 417

seq_name: sp_human:099906
 seq_documentation_block:
 ID 099906 PRELIMINARY; PRT; 417 AA.
 AC 099906;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE D CATEGORY IIIC ANTIGEN.
 GN RHD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66269518; PubMed=8669091;
 RA Beckers E.A., Faas B.H., Ligthart P., Simsek S., Overbeek M.A.,
 RA dem Borne A.E., van Rhenen D.J., der Schoot C.E.;
 RT "Characterization of the hybrid RHD gene leading to the partial D
 RT category IIIC phenotype.";
 RL Transfusion 36:567-574(1996).
 DR EMBL: S82449; AAB37636.2; -.
 DR InterPro: IPR001905; -.
 DR InterPro: IPR002229; -.
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PRINTS: PR00342; RHESUSRHD.
 SQ SEQUENCE 417 AA; 45099 MW; F2A1ECCB5B921260 CRC64;

alignment_scores:
 Quality: 2104.00 Length: 417
 Ratio: 5.058 Gaps: 0
 Percent Similarity: 99.760 Percent Identity: 99.041

alignment_block:
 US-09-600-714-41 x 099906 ..

Align seg 1/1 to: 099906 from: 1 to: 417

1 ATGAGCTCTAGTACCGCGGCTGTCCGGCGCTGCCCTCTGAGGC 50
 1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
 51 CCTAACACTGGAACACAGCTCTCATCTCTCTATTTTTCACCCACT 100
 17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
 101 ATGACGCTTCTTAGAGATCAAAAGGCGCTGTGCAATCCTATCAAGTT 150
 34 YrAspAlaSerLeuGlnAspGlnIleGlyLeuValAlaSerTyrGlnVal 50
 151 GGCAAGATCTGACCGTGAATGGCGGCATGGCTTGGGCTTCCACATC 200
 51 GlnGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuTrpSe 67
 201 GAGTTTCCGAGACACAGCTGAGACAGTGGCTTCAACTTTCATGC 250
 67 rSerPheArgThrHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
 251 TGGCGCTTGGTGTGACAGTGGCAATCTGCTGACAGGCTTCTGAGCCAG 300
 84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
 301 TTCCCTTGGGAGAGTGGTCATACACATGTCATTTGGGCTGGGCGAC 350
 101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
 351 CATGAGTCTTGTGCGTGTGATCTCAGTGGATGCTGTGGGGAAG 400
 117 rMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134
 401 TCAACTTGGCGCAGTTGTGTGTGATGCTGTGAGGTGACAGCTTTTA 450

```

134 aiasnleuaglnleuvalvalmetvalleuvalglvalthrleu 150
451 GGCAACCTGAGGATGTCATCATGTAATATCTTCAACACAGATACACAT 500
151 GlyThrleuArgmetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTTCGACGCTATTTTGGCTGTCTGG 550
167 CAsnMetHisIleTyrValPheAlaIleAlaTyrPheGlyLeuSerValA 184
551 CCTGGTGGCTGGCAAGCTCTACCCGAGGAGGAGGAGATTAATCAG 600
184 IATrPcysLeuProLysProLeuProGluGlyThrGluAspLysAspGln 200
601 ACAGACGATATACCACTTGTCTCCAGTCGCGGGCCCTCTTGTGG 650
201 ThrAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuI 217
651 GATGTTCTGGCCAACTTCAACTCTGTCTGTGTGAGAGTCCAAATGAAA 700
217 pMetThrIleProSerPheAsnSerAlaLeuLeuArgSerProIleGluA 234
701 GGAGAAATGCCGTGTCAACACCTACTATGCTGTAGACGTACGCTGTG 750
234 rGlyAsnAlaValPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCATCTCGAGGTCATCTGCTGGTCACCCCGAAGGAATAGAGA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGluGlyLysIleSer 267
801 GACTATGTCACACTGCGTGTGTGGCAGAGCGCTGCTGGGTACT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValGlyThr 284
851 CGTGCACCTGATCCCTTCTCCGTGGCTGGCATGCTGGCTGTGTG 900
284 eCysHisIleLeuIleProSerProTyrLeuAlaMetValLeuGlyLeuVal 300
901 GGTGGGTCGATCTCGTGGGGGAGGCAAGTACCTGGCGGCTGTGTAA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATTCGCCACAGCTCATGATGGGCTACAACCTCAGCT 1000
317 nArgValLeuGlyIleProHisSerIleMetGlyTyrAsnPheSerL 334
1001 TGTGGGTCTGTGTGAAGATCATCTACATTTGCTGCTGTGTGTGAT 1050
334 euLeuGlyLeuLeuGlyIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTGGAGCGCGCAATGGCATGTGGCTCCAGGTCCTCTCAGCAT 1100
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGCAACTCAGCTTGGCCATGATAGCTCTCAGTCTGCTCTCTCA 1150
367 eGlyLeuLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTGCTCTAAATCTTAAATATGGAAGAACCTCATGAGGCTAAA 1200
384 hGlyLeuLeuLeuAsnLeuLysIleTyrPylAlaProHisGluAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGGTGGAT 1250
401 TyrPheAspAspGlnValPheThrLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9UK19

```

```

seq_documentation_block:
ID Q9UK19 PRELIMINARY; PRT; 417 AA.
AC Q9UK19;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RHD TYPE IIIA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398395; PubMed=9256293;
RA Huang C.H., Chen Y., Reid M.;
RT "Human D(IIa) erythrocytes: Rhd protein is associated with multiple
RT dispersed amino acid variations.";
RL Am. J. Hematol. 55:139-145(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Huang C.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187846; AAF04565.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45273 MW; 491FB1FE1BA37957 CRC64;

```

```

alignment_scores:
Quality: 2094.00 Length: 417
Ratio: 5.058 Gaps: 0
Percent Similarity: 99.281 Percent Identity: 98.801

```

alignment_block:

us-09-600-714-41 x Q9UK19 ..

Align seg 1/1 to: Q9UK19 from: 1 to: 417

```

1 ATGACGTCTAAGTACCGCGGTGTGTCGGGCGCTGCGCTGCGGCG 50
|||||
1 MetSerSerLysTyrProArgSerValArgCysLeuProLeuTyrPAl 17
51 CCTAACACGTGGAAGCAGCTCATCTCTCTCTTATTTTATCCCACT 100
|||||
17 AleuThrLeuGluAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGACGCTTCTTAGAGATCAAAAGGGCTGTGGCATCTATCAAGTT 150
|||||
34 yAspAlaSerLeuGluAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAGATCTGACCGTGATGGCGGCATTTGGCTTGGCTTCTCACTG 200
|||||
51 GlyGlnAspLeuThrValMetAlaLeuGlyLeuGlyPheLeuThrSe 67
201 GAGTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGC 250
|||||
67 rSerPheArgGlyHisSerTyrPheSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGTGTGCAAGTGGCAATCTGTGGAGCGCTCTCTGAGCCAG 300
|||||
84 euAlaLeuGlyValGlnTyrPAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTCTGGGAGAGTGTGTCATCACAGCTTCATATTTGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGTCTTGTGCGTGTGATCTCAAGTGGATCTGTCTTGGGGAAGG 400
|||||
117 rMetSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
401 TCAACTTGGCGCAGTTGCTGTGATGTGCTGTGGAGCTGACACCTTAA 450

```

seq_name: sp_human:Q9UK19

```

|||||
134 alaasleuAlaGlnleuValValMetValleuValGluValThrAlaLeu 150
451 GGCAACCTGGAGATGTCATCAGTAATATCTTCAACACAGACTACACAT 500
151 G|YThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTGTTCGACGCTATTTTGGCTGCTGTGG 550
167 tAsnMetethHisIleTyrValAlaPheAlaIleTyrPheGlyLeuSerValA 184
551 CCTGTGCTGCCCAAGCCTCTACCCGAGGACGAGAGATAAAGATCAG 600
184 lArpCysLeuProLysProLeuProGluGlyThrGluAspLysAspGln 200
601 ACAGAACGATACCAAGTTGTGTCGACCTGCTGGGCGCCCTCTCTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGCTGTGAGAAATCCAAATCGAA 700
217 pMetPheTrpProSerValAlaSerAlaLeuLeuArgSerProIleGluA 234
701 GGAAGATGCCGTGTTCACACACTACTAGTGTAGCAGTACGCTGTG 750
234 rGlyAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGGCATCTGAGGCTCATCTGCTGCTACGCCCAAGGAGATCAGCAA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerLys 267
801 GACTATGTGCACAGTGCCTGTTCGACGAGGCGTGGCTGTGGGTACT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyValAlaValGlyThrS 284
851 CGTGCACCTGATCCCTTCTCCGTGGCTTGCCTGCAATGCTGCTGTG 900
284 eRcysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCGTGGGGGAGCCAAATGACTGCGGGGCTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
951 CCGAGTGTGGGATTTCCCAACAGTCCATGATGGGCTTACAACTTCACT 1000
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
1001 TGTGGGCTCTCTGGAGAGATCATCATATGTTGCTGCTGCTGTGAT 1050
334 euleuGlyLeuLeuGlyGlyIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTCGAGCCGGCAATGGCATGTTGGCTTCCAGGCTCCCTCAGCAT 1100
351 ThrValGlyAlaGlyAsnGlyMetCileGlyPheGlnValLeuLeuSerL 367
1101 TGGGGAACCTGAGTTGGCCATCTGATAGTCTGACGCTGCTGCTCTGA 1150
367 eGlyGlyLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTGCTCTTAATCTTAATATGGAAGAAGCCTCATGAGGCTTAA 1200
384 hrGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlyAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGGATTT 1250
401 TyrPheAspAspGlnValAlaPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9HCC3

```

```

seq_documentation_block:
ID Q9HCC3 PRELIMINARY; PRT; 417 AA.
AC Q9HCC3:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE RH BLOOD D GROUP ANTIGEN POLYPEPTIDE.
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
ON (1)
RP SEQUENCE FROM N.A.
RA Kashiwase K., Uchikawa M., Ishikawa Y.;
RT "Rh blood D group antigen polypeptide, RHDK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049754; BAB16598.1; -
SO SEQUENCE 417 AA; 45136 MW; 2CA89E3E716087C8 CRC64;

```

```

alignment_scores:
Quality: 2094.00 Length: 417
Ratio: 5.046 Gaps: 0
Percent Similarity: 99.520 Percent Identity: 98.321

```

alignment_block:

US-09-600-714-41 x Q9HCC3

Align seg 1/1 to: Q9HCC3 from: 1 to: 417

```

1 ATGAGCTCTAAGTACCGCGGTCTGTCCGGCGTCCCTGCGCTGTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTrpAl 17
51 CCTTACACTGGAGACACTCTCATTTCTCTTATTTTATTTTACCACACT 100
|||||
17 aleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGACGGTTCCTTAGAGATCAAAAGGGCGTCGTGCATCTTACAGTT 150
|||||
34 yrasPalaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAAGATCGACCGTGAATGGCGCCATTTGGCTTGGCTTCTCACCTC 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaLeuGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGGAGACACAGCTGGAGCACTGTGGCTTCAACCTTTTCATGC 250
|||||
67 rAsnPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGGCGTGTGTGTCAGTGGGCAATCGCTGGAGGCTTCCTGAGGCAG 300
|||||
84 euAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 300
301 TTCCCTTCTGGGAAGTGTGCATCACACTGTTTCAGTATTCGGCTGGCCAC 350
|||||
101 PheTrpProGlyLysValValIleThrLeuPheSerIleThrGlyLeuAlaThr 367
351 CATGAGTCTTTTGTGGTGTGATCTGATCGATGAGATGCTGTGTGGGAAGG 400
|||||
117 rMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 417
401 TCACTTGGCGCCAGTGTGGTGTGATGTGCTGTGGAGGAGTTCATTA 450
|||||
134 alaAsnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCAACCTGGAGATGTCATCAGTAATATCTTCAACACAGACTACACAT 500
|||||
151 G|YThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGACATCTACGTGTTCGACGCTATTTTGGCTGCTGTGG 550
|||||

```

```

167  tasmethethisletyvalphealalatyphneglyleuservala 184
551  cctgtgtcctgcaaaagccttacccgagagagagataaagatcag 600
184  latpccysleuprolysproleuproglythngluaspyspoin 200
601  acagcaacgataccagattgtctgcccattgctggcgccctctctgtg 650
201  thratathrlleproserleuserlametleuglyalaleupheletr 217
651  gatgtcttgcccaagttccaactgctgctgctgagagatccaatcgaaa 700
217  pmetheptrproserpheasnservalaleuuarserproleglua 234
701  ggaacaaatgccgtgtcaaacctctactgtgctgagatcagctggtg 750
234  tglvasasnlavalpheasnthyrytyralavalalavalaservala 250
751  acagccatctcagaggtcatccctggtcaccgccccaagagagatcagca 800
251  thralalleserglyserleuserlametleuglyalaleuvalleuval 267
801  gacttatgtgcacagctgctgctggtggcagagagcgtgctggtacct 850
267  sthrytyvalhisseralavalaleualaglyalavalavalaglythr 284
851  cgtgcacactgattccctctcgcgtggtgagatgctgctggtgctggtg 900
284  ecysnhsleuileproserprotripleuvalametalaleuvalleuval 300
901  gctggcctgattctccctgctggggagagcacaagatcctgcccgggtgtgtaa 950
301  alaglyleuileservalaglyalalystryleuproglycyscysas 317
951  ccgagctgctggggattcccccacagctccatctatgggctacaaactcagct 1000
317  narvalleuglyleuprohisserleuemetglytyrasnpheletr 334
1001  tgcctggctgctgctggagagatcatctacatgtgctgctgctgctgctgct 1050
334  euleuglyleuileuglylunilleuvalleuvalleuvalleuval 350
1051  accgctgagagccgcaatggcattgctgctgctgctgctgctgctgctgct 1100
351  thrvalaglyalaglyasnlymetileglypheglnvalleuvalleuval 367
1101  tggggaactcagctgctgctgctgctgctgctgctgctgctgctgctgctg 1150
367  eelyleuileuserleuvalaleuvalleuvalleuvalleuvalleuval 384
1151  caggttgctgcttaaatcttaaaatggaagacacccatgagagctaa 1200
384  hcglyleuileuasnleuvalleuvalleuvalleuvalleuvalleuval 400
1201  tattttgatgacaaagtttctggaagtttccatgttgctgttgatt 1250
401  tyrpheasaspglvalpheptrpyspheprohisleuvalavalaglyph 417
1251  T 1251
417  e 417

```

seq_name: sp_human:Q9H248

seq_documentation_block:

ID: Q9H248 PREDIMINARY: PRT: 417 AA.

AC: Q9H248:

DT: 01-MAR-2001 (TREMBLERL. 16, Created)

DT: 01-MAR-2001 (TREMBLERL. 16, Last sequence update)

DT: 01-MAR-2001 (TREMBLERL. 16, Last annotation update)

DE: RHESUS D CATEGORY VI TYPE IV PROTEIN.

GN: RHD.

OS: Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Esteban R., Nogues N., Montero R., Hurtado M., Maroto S., Ribera A.;
RT "Characterization of a new RHD category VI (type IV) allele,
RT predominant in the Spanish population.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF12679; AAC40872.1; -; 03FD058960E41D89 CRC64;
SQ SEQUENCE 417 AA; 45293 MW; 03FD058960E41D89 CRC64;

```

alignment_scores:

Quality:	2043.00	Length:	417
Ratio:	4.971	Gaps:	0
Percent Similarity:	98.561	Percent Identity:	95.923

alignment_block:

US-09-600-714-41 x Q9H248

Align seg 1/1 to: Q9H248 from: 1 to: 417

```

1  ATGAGCTCTTAAGTACCGCGGCTGTGTCGGCGCTGCTGCCCTCTGGGC 50
|||||
1  MetSerSerlytyrProArgSerValAlrArgCysLeuProLeuTrpAl 17
51  CCTACACTGGAAGACGCTCATCTCTCTCTATTTTACCACCT 100
|||||
17  AleuthrleugluAlaAlaLeuileLeuLeuphethyrpethrhist 34
101  ATGAGCTTCTCTTAAGAGATCAAAAGGGCTGCTGCACTCATCAAGTT 150
|||||
34  YrAspAlaSerleuLunspGlnlysglyleuValAlaSerlyrGlnAl 50
151  GGCCAGATCTGACCGTGAATGGCGGCATTGGCTGGCTTCTTCACCTC 200
51  GlyGlnAspLeuThrValMetAlaAlaileglyleuglypheleuThrse 67
201  GAGCTTCCGGAGACACAGCTGAGAGAGTGGCTTCAACCTTCATATGC 250
67  rserphearlgarGhisserTrpSerSerValAlaPheasnleuphemel 84
251  TGGCGCTTGTGTGTCAGTGGGCAATCTGCTGAGAGGCTTCTCGAGCCAG 300
84  eulAlaLeuglyValGlnTrpAlaileLeuLeuAspGlyPheLeuSerGln 100
301  TTCCTTCTGGGAAGGTGTCATCACACTGTCAGTATTGGCTGGCCAC 350
101  PheProSerGlyLyValValAlleThrLeupheSerileArgLeuAlaTh 117
351  CATGAGTGGTGTGCTGGGCTGATCTCAAGTGAATGCTGCTGGGGAAGG 400
117  rmetserAlaMetSerValLeuileSerAlaGlyAlaValleuglylysv 134
401  TCNACTTGGCGCAGTTGTGTGATGTGCTGCTGGTGAAGTGAACCTTTA 450
134  AlAsnLeuAlaGlnLeuValValMetValleuValGluValThrAlaLeu 150
451  GGCAACCTGAGAGATGTCATCAGTAATATCTTCAACACAGATACCAAT 500
151  GlyThrLeuAlrGmetValileSerAsnIlePheAsnIleAspTyrlHisMe 167
501  GAACATGATGACATCTACGTTGTGCGAGCCATTTTGGGCTGCTGGG 550
167  tasnleuArghisphetyrValAlpheAlaAlaTyrlpheglyleuThrVal 184
551  CCTGTGCTGCTGCAAAAGCCTTACCCGAGGAGCAAGAGATAAAGATCAG 600
184  latpccysleurolysproleuprolysglythngluaspyspoin 200
601  ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGCGCCCTTTCTTGTG 650
|||||

```

```

201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTTCGGCCAGTTTCACATCGTCTGCTGAGAGTCCATTCGAAA 700
|||||
217 PheThrProSerValAsnSerAlaLeuLeuArgSerProIleGlnA 234
|||||
701 GGAAGAAATGCGGTGTTCACACCTACTAGTGTAGACAGTACGGGTGTG 750
|||||
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
|||||
751 ACAGCCATCTCAGAGGTTCATCTTGCTGCTACCCCAAGGAGATCAGAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
|||||
801 GACTTATGTGCACAGTGGGTGTGGAGAGAGCGTGGCTGGTACT 850
|||||
267 tThrTyrAlaHisSerAlaValLeuAlaGlyAlaValAlaValGlyThrS 284
|||||
851 CGTGTACCTGATCCCTTCCTGCTGGCTTGGCATGTGCTGGTGTG 900
|||||
284 eRcYHisIleLeuIleProSerProThrPleuAlaMetValLeuGlyLeuVal 300
|||||
901 GCTGGGCTGATCTCCGCTGGGGGAGCCAGTACTGCCGGGGGTGTGTA 950
|||||
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyCysCysAs 317
|||||
951 CCGAGTGGTGGGATCCCGACAGTCCATCATGGGCTACAACTTCAGCT 1000
|||||
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
|||||
1001 TGCCTGGCTCTCTGGAGAGATCATCTACATTTGCTGCTGCTGTGAT 1050
|||||
334 euleuGlyLeuLeuGlyIleIleIleTyrIleValLeuLeuValLeuAsp 350
|||||
1051 ACCGTGGAGCCGGCAATGCGCATGTTGGCTTCCAGGCTCTCCAGCAT 1100
|||||
351 ThrValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerL 367
|||||
1101 TGGGGAATCAGCTTGAGCATGATAGTACGTCTGACGCTGCTGCCCTGA 1150
|||||
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
|||||
1151 CAGGTTTCTCTTAATCTTAATATATGAAAGACCTCATGAGGCTAAA 1200
|||||
384 hngIyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
|||||
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAT 1250
|||||
401 TyrPheAspAspGlnValIleThrLysPheProHisLeuAlaValGlyPh 417
|||||
1251 T 1251
|
417 e 417

```

seq_name: sp_human:Q9NPX0

seq_documentation_block:

ID Q9NPX0 PRELIMINARY: PRT: 417 AA.

AC Q9NPX0:

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE RH BLOOD GROUP ANTIGEN RHD.

GN RHIVB(J).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hyodo H., Ishikawa Y., Tsuneyama H., Kashiwase K., Toyoda C.,

RA Uchikawa M., Akaza T., Fujii T., Kozuma S., Takekani Y., Juji T.,

RT "New Rhivb identified in Japanese."

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB037270; BAA90298.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45429 MW; 647D44A7678CB498 CRC64;

alignment_scores:
  Quality: 2039.00      Length: 417
  Ratio: 5.022          Gaps: 0
  Percent Similarity: 97.362      Percent Identity: 97.122

alignment_block:
  US-09-600-714-41 x Q9NPX0

Align seg 1/1 to: Q9NPX0 from: 1 to: 417

1 ATGAGCTCTAAGTACCGCGGCTGTCCGGCGCTGCCCTGGCCCTGGGC 50
|||||
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuTrpAl 17
|||||
51 CCTAACCTGGAGACAGCTCTCATTTCTCTTATTTTTCACCACT 100
|||||
17 AleuThrLeuGlnAlaAlaLeuIleLeuPheTyrPhePheThrHisT 34
|||||
101 ATGACGCTTCCTTAGAGATCAAAAGGGGCTCGTGGATCCTATACAGTT 150
|||||
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
|||||
151 GGCCAGATCTGACCGTATGAGCGGCCATTTGGCTTGGCTTCACCTC 200
|||||
51 GlyLinspleuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrS 67
|||||
201 GAGTTTCGGAGACACAGCTGAGCAGTGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
|||||
251 TGGGCTTGTGTGTCAGTGGGCAATCCTGCTGGAGCGCTTCCATGAGCAG 300
|||||
84 euaAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
|||||
301 TTCCCTTCTGGGAGGTGTCATCACACTGTTCAATTCGGCTGGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
|||||
351 CATGAGTCTTGTGCTGCTGATCTCAGTGGATGCTGTCTTGGGGAAAG 400
|||||
117 rMetSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
|||||
401 TCAACTTGGCGCAGTGTGCTGATGGTGGTGGTGGAGGTGACAGCTTTA 450
|||||
134 aLAsnLeuAlaGlnLeuValValaMetValaLeuValaGlnValaThrAlaLeu 150
|||||
451 GGCACTGAGAGATGTCATCAGTAATATCTTCAACACAGACTACCAAT 500
|||||
151 GlyAsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
|||||
501 GAACATGATGACATCTACGTTGCGCAGCCTATTGGGCTGTCTGTGG 550
|||||
167 tAsnMetMetHisIleTyrValaPheAlaAlaTyrPheGlyLeuSerVala 184
|||||
551 CCTGTGGCTGCGCAAGCTCTACCCGAGGAGACGAGAGATAAGATCAG 600
|||||
184 lArgPcysLeuProLysProLeuProGlnGlyTrpGlnAspLysAspGln 200
|||||
601 ACAGCAACGATACCACTTGTCTGCATGCTGTGGCGCCCTTCTTGTG 650
|||||
201 ThrIleThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
|||||
651 GATGTTTCGGCCAGTTTCACATCTGCTGCTGCTGAGAGTCCATTCGAAA 700
|||||

```



```

217 pNebHeTrrProSerPheAsnSerAlaLeuLyrSerProIleGluA 234
701 GGAAGATGCGCTGTTCACACCTACTATGCTGACAGTCAGCGTGG 750
|||||
234 rGLyAsnAlaValAlpheAsnThrTyrTAlaValAlaValSerValAl 250
|||||
751 ACAGCCATCTCAGGGTCATCTCTGGCTCACCCCAAGGAGATCAGCAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnGlyIleSerly 267
|||||
801 GACTATGTCACAGTCAGCGCTGTGGCAGAGCGCTGGCTGGTACCT 850
|||||
267 sThrTyrValHisSerAlaValLeuAlaGlyIlaValaValGlyThrS 284
|||||
851 CGTGTACCTGATCCCTCTCCGTGGCTGACCATGCTGGCTGGCTGG 900
|||||
284 eTcYHisIleuIleProSerProTrrPleuAlaMetValLeuGlyLeuVal 300
|||||
901 GCTGGGCTGATCTCCGTGGGAGCCCAAGTACCTGCCGGGTGTGTTAA 950
|||||
301 AlAGlyLeuIleSerValGlyIaValysTyrLeuProValIcysCysAs 317
|||||
951 CCGAGTCGTGGGGATCCCCACAGCTCCATCATGGGCTCAACTTCAGCT 1000
|||||
317 nArgValLeuGlyIleHisIleSerValMetHisSerIlePheSerL 334
|||||
1001 TCCTGGTCTGCTTGAGAGATCATCTACATGCTGCTGCTGCTGAT 1050
|||||
334 euLeuGlyLeuLeuGlyIleuIleThrTyrIleValLeuLeuValLeuHis 350
|||||
1051 ACCGTGCGAGCCGCGCAATGCGATGATGGCTTCAGGCTCCCTCAGCAT 1100
|||||
351 ThrValTrrPasnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerL 367
|||||
1101 TGGGGAACCTCAGCTGGCCATGCTATAGCTCCTCAGCTGCTGCTCCTTA 1150
|||||
367 eGlyLeuLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
|||||
1151 CAGGTTTGCTCTTAATCTTAATATATGGAAGCACCTCATGAGGCTAAA 1200
|||||
384 hGcIyLeuLeuLeuAsnLeuysIleTrrPlyAlaProHisValAlaLys 400
|||||
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGGTGAAT 1250
|||||
401 TyrPheAspAspGlnValAlpheTrrPlySPheProHisLeuAlaValGlyh 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9UDZ5
seq_documentation_block:
ID Q9UDZ5 PRELIMINARY; PRT; 417 AA.
AC Q9UDZ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RHESUS D CATEGORY VI TYPE III PROTEIN.
CN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=98158656; PubMed=9490704;
RA Wagner F.F., Gassner C., Mueller T.H., Schoenitzer D., Schunter F.,
RT "Three molecular structures cause Rhesus D category VI phenotypes with
distinct immunohematologic features.";
RL Blood 91:2157-2168(1998).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=94245182; PubMed=8188244;
RA Cherif-Zahar B., Le Van Kim C., Rouillac C., Raynal V., Cartlon J.,
RA Collin Y.;
RT "Organization of the gene (RHCE) encoding the human blood group Rhcbe
RT antigens and characterization of the promoter region.";
RL Genomics 19:68-74(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=9306356; PubMed=1438298;
RA Le Van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
RA Cartlon J.P., Collin Y.;
RT "Molecular cloning and primary structure of the human blood group Rhd
RT polypeptide.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
DR EMBL; Z97026; CAB09722.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45247 MW; 1887A3C519149F8F CRC64;

alignment_scores:
Quality: 2033.00 Length: 417
Ratio: 4.959 Gaps: 0
Percent Similarity: 98.321 Percent Identity: 95.444

alignment_block:
US-09-600-714-41 x Q9UDZ5
Align seg 1/1 to: Q9UDZ5 from: 1 to: 417

1 ATGAGCTCTAAGTACCCGGCTGTCTCCGCGCTGCTGCCCTGCCCTGGGC 50
|||||
1 MetSerSerLySTyrProArgSerValArgGlyLeuProLeuTrrPAl 17
|||||
51 CCTACACCTGGAAGCAGCTCTCATCTCTCTCTTATTTTACCACCT 100
|||||
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHist 34
|||||
101 ATGAGCCTCTTAGAGATCAAAAGGGCTGCTGGCACTCATCAAGT 150
|||||
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
|||||
151 GGCCAAAGATCTGACCGTATGGCGGCATTTGGCTTGCTCCACCTC 200
|||||
51 GlyGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
|||||
201 GAGTTTCCGAGACACACAGCTGAGCAGTGTGGCTTCAACCTTTCATGC 250
|||||
67 rSerPheArgGlyHisSerTrrPrrSerSerValAlaPheAsnLeuPheMetL 84
|||||
251 TGGCGCTTGTGGTGGAGTGGGCAATCCCTGCGAGCGGCTCTCGAGCCAG 300
|||||
84 euAlaLeuGlyIaGlnTrrPAlaIleLeuLeuAspLysPheLeuSerGln 100
|||||
301 TTCCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTCGGCTGCCAC 350
|||||
101 PheProSerGlyLysValValIleThrIleuPheSerIleArgLeuAlaThr 117
|||||
351 CATGAGTCTTGTGGTCTGCTCATCTCACTGATGCTGCTTGGGGAAG 400
|||||
117 TrMetSerAlaMetSerValLeuIleSerAlaGlyIaValaValLeuGlyLysV 134
|||||
401 TCAACTGGCGCAGTGTGTGATGATGCTGCTGGAGTGCAGACTTAA 450
|||||
134 AlAsnLeuAlaGlnLeuValValMetValLeuValGlyIaValThrAlaLeu 150
|||||
451 GGCAACCTGAGATGATGTCATCATATATCTTCAACACAGACTACCAT 500

```

```

|||||
151 GlyThrLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMet 167
501 GAACATGATGACATCTAGCTTGGACCATATTTGGCGTCTGCTG 550
167 TAsnLeuArgHisPheTyrValPheAlaIleTyrPheGlyLeuThrValAla 184
551 CCTGGTGCCTGCCAAAGCCTTACCCGAGGAGGAGATGAATGAC 600
184 IATrCysLeuProIlySerProIlySerGlyTyrGlnAspAsnAspGln 200
601 ACAGCAGCATACCCAGTTTGTCTGCCATGCTGGCGCCCTTCTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCCAAGTTTCACTGCTGCTGCTGAGAAATCCAAATCGAAA 700
217 PheTrpPheTrpProSerValAsnSerAlaLeuLeuArgSerProIleGlnA 234
701 GGAAGAAATGCCTGTTCACACCTACTATGCTGTAGACATGACGCGTGTG 750
234 rGlyAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValAla 250
751 ACAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGAGATGACGAA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgGlyIleSerMet 267
801 GACTATATGTCACAGTGGCGGTGTGGCAGAGAGCGTGGTGGTACTCT 850
267 TThrTyrValHisSerAlaValLeuAlaGlyIleValAlaValAlaGlyThrS 284
851 CGTGTACCTGATCCCTTCTCCGTGGCTGACATGCTGCTGCTGTGTG 900
284 eTyrHisIleLeuIleProSerProTrpLeuAlaMetValLeuVal 300
901 GCTGGCGTGTATCTCCGTGGGAGGAGCAATGACTCCGCGGCTGTGTAA 950
301 AlaGlyLeuIleSerIleGlyIleValAlaGlyCysLeuProGlyCysCysAs 317
951 CCGAGTGGCGGGGATTCGCCACAGCTCATGCGGCTTACAACTTCAGCT 1000
317 nArgValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheSerL 334
1001 TCGTGGTCTGCTTGGAGATCATCTACATGCTGCTGCTGCTGCTGAT 1050
334 euleuGlyLeuLeuGlyIleIleTyrIleValLeuLeuValLeuAsp 350
1051 ACCGTGCGAGCCGCAATGGCATGATGGCTTCCAGGTCTCTCAGCAT 1100
351 ThrValIleGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGAGACTCAGCTGGCCATGCTGATAGCTTCCAGTCTGCTGCTCCGTA 1150
367 eGlyIleLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTTGCCTTAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
384 hGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
401 TyrPheAspAspGlnValAlaPheTrpLysPheProHisLeuAlaValAlaGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human:Q9UK74
seq_documentation_block:
ID Q9UK74; PRELIMINARY; PRT; 417 AA.
AC Q9UK74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE RH BLOOD GROUP ANTIGEN EVANS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RX MEDLINE=96404415; PubMed=8808597;
RA Huang C.H., Chen Y., Reid M., Ghosh S.;
RT "Genetic recombination at the human RH locus: a family study of the
RT red-cell Evans phenotype reveals a transfer of exons 2-6 from the RHD
RT to the RHCE gene."
RL Am. J. Hum. Genet. 59:825-833(1996).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RA Huang C.H.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF17938; AAD55773.1; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp. 1.
DR PRINTS; PR00342; RHESUSRHD.
DR ProDom; PD000664; -.
SQ SEQUENCE 417 AA; 45445 MW; D8AA0B148ADCT3F7 CRC64;

alignment_scores:
Quality: 2016.00 Length: 417
Ratio: 4.990 Gaps: 0
Percent Similarity: 96.882 Percent Identity: 96.403

alignment_block:
US-09-600-714-41 x Q9UK74

Align seg 1/1 to: Q9UK74 from: 1 to: 417

1 ATGAGCTCTAAGTACCGCGGCTGTGTCCGGCGCTGCGCCCTCTGGGC 50
1 MetSerSerLysTyrProArgSerValArgArgCysLeuProLeuCysAl 17
51 CCTACACTGAGACAGCTCTCATCTCCCTCTCATTTTACCACCT 100
17 AleuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisT 34
101 ATGACGCTTCCCTTAGAGGATCAAAAGGGGCTCGTGCAATCCTATCAAGTT 150
34 YrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GAGCAAGATCTGACCGGTGATGGCGCATTTGGCTTGGCTTCTCTACCTC 200
51 GlyGlnAspLeuThrValMetAlaIleLeuGlyLeuGlyPheLeuThrSe 67
201 GAGTTCCGGAGACACAGCTGAGAGCTGAGCGCTTCAACCTTCATGAC 250
67 rSerPheArgArgHisSerTrpSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGTGTGACAGTGGCAATCTGCTGACAGCGCTTCCAGCCAG 300
84 eAlaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCTGGGAAGGTGTCATCACACTGTTCAAGTATTCGGTGGCCAC 350
101 PheProSerGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
351 CATGAGTCTTGTGCGGTGATGATCTCAGTGAGTGTGCTGTGGGAAG 400
117 rMetSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134

```

```

401 TCAACTGGCGCAGTGGTGGTGAATGTCGTGGAGCTGACAGCTTAA 450
134 aAlaenLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCTGGAGATGTCATCAGTAATATCTTCAACACAGACTACCAT 500
151 GlysAlaLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTTCGCGAGCACTATTTTGGCTGTGG 550
167 tAsmMetHisIleTyrValPheAlaIleTyrPheGlyLeuSerValA 184
551 CCGTGCGCTGCCAAAGCCTTACCAGGAGAACGAGATAAAGATCAG 600
184 latrPcysLeuProIysProIysProIysGlnGlyThrGlnAspLysaspI 200
601 ACAGCAACGATACCAAGTTTGTCTGCCATGCTGGGCGCCCTTCTTG 650
201 ThrIatThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeu 217
651 GATGTTTGGCCAAAGTTTCAACTGCTGCTGCTGAGAGATCAATCGAAA 700
217 pMetPheThrProSerPheAsnSerAlaLeuLeuArgSerProIleGln 234
701 GGAAGATGCGGTGTTCACACCTACTATGCTGTAGACAGTCAGCTGTG 750
234 rGlyAsnAlaValPheAsnThrTyrTyrAlaValAlaValSerValVal 250
751 ACAGCAACCTGAGGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerLy 267
801 GACTATGTCGACAGTGGCTGTGGTGGAGAGCGCTGGCTGGTGGTACCT 850
267 sThrTyrValHisSerAlaValLeuAlaGlyGlyAlaValAlaGlyThr 284
851 CGGTGACCTGATCCCTTCTCCGTGGCTGCCATGCTGCTGCTGCTGTG 900
284 eRcysHisLeuIleProSerProIlePheAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCCGTGGGAGGAGCAAGTACCTGCCGGGCTGTGTGTA 950
301 AlaGlyLeuIleSerValGlyGlyAlaLysTyrLeuProValLysCysAs 317
951 CCGAGTGTGGGATTCACACAGCTCCATGAGGCTACACACTCAGCT 1000
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGTGGGTCTGCTGGAGATCATCTACATTTGCTGCTGCTGCTGCTGAT 1050
334 euLeuGlyLeuLeuGlyIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTGGAGCGCGCATATGATGATTTGGCTTCCAGTCTCTCAGCAT 1100
351 ThrValTyrAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGAACTGAGCTGGCCATCGATAGCTCTGCTGCTGCTGCTGCTGA 1150
367 eGlyIleLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTTCTCTTAATCTTAATATATGAAAGCAGCTCATGAGGCTTAA 1200
384 hnglyLeuLeuLeuAsnLeuLysIleTyrLysAlaProHisValAlaLys 400
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
401 TyrPheAspAspGlnValPheThrLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
seq_name: sp_human: Q9UPN0

```

```

seq_documentation_block:
ID Q9UPN0 PRELIMINARY; PRT; 417 AA.
AC Q9UPN0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2001 (TREMBlrel. 16, Last annotation update)
DE RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.
GN RHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashlase K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variant in
RT Japanese.";
RN [2]
RP SEQUENCE FROM N.A.
RA Kashlase K., Uchikawa M., Ishikawa Y.;
RT "Rh blood CE group antigen polypeptide, RHEK.";
RN [3]
RP Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030388; BA82627.1; -
DR EMBL; AB049753; BAB16597.1; -
DR InterPro; IPR001905; -
DR InterPro; IPR002229; -
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
SQ SEQUENCE 417 AA; 45604 MW; 61E4BB39ABA23CC CRC64;

```

```

alignment_scores:
Quality: 1963.00 Length: 417
Ratio: 4.920 Gaps: 0
Percent Similarity: 95.683 Percent Identity: 93.285

```

```

alignment_block:
US-09-600-714-41 x Q9UPN0

```

```

Align seg 1/1 to: Q9UPN0 from: 1 to: 417

```

```

1 ATGAGCTCTAAGTACCCGCGGTCTGTCGCGGCTGCTGCCCTCTGGGC 50
1 MetSerSerLysTyrProArgSerValAlaArgCysLeuProLeuTyrAl 17
51 CCTAAGCTGGAAGCAGCTCTCATTTCTCTCTTATTTTATCCCACT 100
17 aLeuThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHis 34
101 ATGAGCTCTCTTACAGATCAAAAGGGGCTGCGCATCCATCAAGATT 150
34 yAspAlaSerLeuGlnLysGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCTAAGATCTGACCGTATGCGGCGCATTTGGCTTGGCTTCCACCTC 200
51 GlysGlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSe 67
201 GAGTTTCCGAGACACACAGCTGAGAGAGTGTGGCTTCAACCTCTTCATGC 250
67 rSerPheArgAlaGHisSerTyrSerSerValAlaIleAsnLeuPheMetL 84
251 TGGCGCTGTGTGTCAGTGGGCAATCCGTCGTGGAGGCTCTCTAGCCAG 300
84 euAlaLeuGlyValGlnTyrPheAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCCTTCTGGAAGGTGTCATCACACTGTTCATATTCGCTGCGCAC 350
101 PheProSerGlyLysValValIleThrIlePheSerIleAlaGlyLeuAlaTh 117
351 CATGAGTCTTTGTGCGGTGATCTCAGTGGATGCTCTTGGGGAAG 400

```

```

117 mterSerAlaLeuSerValLeuIleSerValAspAlaValLeuGlyLysV 134
401 TCAACTGGCCGCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 450
134 aIaenLeuAlaGlnLeuValValMetValLeuValGlnValThrAlaLeu 150
451 GGCAACCGGAGATGTCATGATTAATATCTTCAACACAGACATACACAT 500
151 GlysLeuLeuArgMetValIleSerAsnIlePheSerThrPheSerThr 167
501 GAACATGATGCACATCTACGTGGTGGAGCCATATTTGGGCTGTCTGG 550
167 tAsnLeuArgHisPheThrValPheAlaAlaValPheGlyLeuThrVal 184
551 CCTGGTGGCTGGCAAGCCCTACCCGAGAGAGAGAGAGAGAGAGAGAGAG 600
184 larrPcysLeuProLysProLeuProLysGlyThrGlnAspAsnAspGln 200
601 ACAGCAACGATACCCAGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTCTGGCCCAAGTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
217 pMetPheThrProSerValAsnSerProLeuLeuArgSerProIleGln 234
701 GGAGAAATGCCGTGTCAACACCTACTATGCTGACAGAGTACAGTGGTG 750
234 rGlyAsnAlaMetPheAsnThrTyrrAlaLeuAlaValSerValVal 250
751 ACAGCAACCTCAGGGTCATCCTTGGCTACCCCAAGGAGAGATACCA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTAATGTGCACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 850
267 tThrTyrrValHisSerAlaValLeuAlaGlyValAlaValAlaGlyThr 284
851 CGTGCACCTGATCCCTCTCCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
284 erGySHisLeuIleProSerProThrPheAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGAGGAGCAAGTACCTGCCGGGGTGTGTA 950
301 AlaGlyLeuIleSerIleGlyAlaLysCysLeuProValCysCysAs 317
951 CCGAGTGTGGGGATTCGCCACAGCTCATCATGGGCTACAACTTCAGCT 1000
317 nArgValLeuGlyIleHisIleSerValMetHisSerIlePheSerL 334
1001 TGTGGGCTGCTGGAGAGATCATCATGATGTCGTGGTGGTGGTGGT 1050
334 euLeuGlyLeuLeuGlyIleThrTyrrIleValLeuLeuValLeuHis 350
1051 ACCGTGGAGCGGCAATGGCATGATGGCTTCAGAGTCTCTCAGCAT 1100
351 ThrValIrrPaangLysangLysMetIleGlyPheGlnValIleLeuSer 367
1101 TGGGGAACTCAGCTTGGCCATCGATAGCTTCACGCTGTGTCTCGGA 1150
367 eGlyGlnLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeu 384
1151 CAGGTTGGCTCTAAATCTTAAATATGGAAGACCTCATGAGAGCTAAA 1200
384 hrGlyLeuLeuLeuAsnLeuLysIleThrPlySalProHisValAlaLys 400
1201 TATTTTGAAGCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
401 TyrPheAspArgGlnValPheThrPlyPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_name: sp_human:Q9UEC3
seq_documentation_block:
ID Q9UEC3 PRELIMINARY; PRT; 417 AA.
AC Q9UEC3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RH BLOOD CE GROUP ANTIGEN POLYPEPTIDE.
GN RHCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashikawa K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variants
RT (RHEPM and RHEKH) in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018644; BAA33927.1; -.
DR InterPro; IPR000885; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
DR ProDom; PD002078; -.
DR SEQUENCE 417 AA; 45529 MW; 3B072CA1544F48C0 CRC64;

alignment_scores:
Quality: 1940.00 Length: 417
Ratio: 4.887 Gaps: 0
Percent Similarity: 95.204 Percent Identity: 92.086

alignment_block:
US-09-600-714-41 x Q9UEC3 ..
Align seg 1/1 to: Q9UEC3 from: 1 to: 417

1 ATGAGCTTAAGTACCCGGGCTGTCCGGGCTGCCCTGCCCCCTGGGC 50
1 MetSerLeuLysTyrrProArgSerValArgArgCysLeuProLeuThrPAl 17
51 CCTAACCTGGAGACAGCTCATCTCCCTCTCTATTTTTCACCACT 100
17 aLeuThrLeuGlnAlaIleLeuIleLeuLeuPheTyrrPheThrHisT 34
101 ATGACGCTTCCTTAGAGATCAAAAGGGCTCGTGCATCCTATCAAGTT 150
34 yrrAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrrGlnVal 50
151 GGCCAAAGATTCAGCGTGATGGCGGCAATGGCTTGGGCTTCTCACCTC 200
51 GlysGlnAspLeuThrValMetAlaAlaLeuGlyLeuGlyPheLeuThrse 67
201 GAGTTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAACCTTTCATGC 250
67 tAsnPheArgThrHisSerThrPheSerSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGGTGTGACATGGGCAATCCTGTGAGAGGCTTCTGAGGCAG 300
84 euAlaLeuGlyValGlnThrPAlaIleLeuLeuAspGlyPheLeuSerGln 100
301 TTCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTTGGCTGGCCAC 350
101 PheProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThr 117
351 CATGAGGCTTGTGCGGTGATCGATCGATGAGATCGTGTGGGAGAG 400
117 mterSerAlaLeuSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134

```

seq_name: sp_human:Q9UEC2

```

401 TCAACTGGCGCAGTTGGTGTGATGGTGTGCTGGAGAGTACAGCTTTA 450
134 aLaAsnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150
451 GGCACCGCTGAGGATGTCATGATGTAATATCTTCAACACAGCTACCAT 500
151 GLyThrLeuAlaGMeValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCACATCTACGTGTTCGACAGCTAATTTGGCTGTCTGG 550
167 tAsnLeuAlaGHisPheTyrValPheAlaIaLalTyrPheGlyLeuThrValA 184
551 CCTGGCTGCTGCCAAAGCTTACCCGAGGAGACGAGATGAAGATCAG 600
184 latrPcysLeuProLysProLeuProLysGlyThrGluAspAsnAspGln 200
601 ACAGAACGATACCCAGTTGTGTCGCTGCTGGCGGCCCTCTCTTGTG 650
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATATTGTGGCCAAAGTTTCAACTGTGCTGTGCTGAGAGTCCAAATCGAA 700
217 pMetPheThrProSerValAsnSerProLeuLeuArgSerProIleGluA 234
701 GGAAGATGCGGTGTTCACACCTACTAGTGTGAGCAGTACGCGGTGG 750
234 rGlyAsnAlaValPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCCATCTCAGGGTTCATCTTGGCTACCCCGCAGGAGATCAGCA 800
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTATGTGCACAGTGGGTGTGGTGGCAGAGCGGTGGCTGGGTACT 850
267 tThrTyrValHisSerAlaValLeuAlaGlyGlyAlaValAlaGlyThrS 284
851 CGTGTACCTGATCCCTTCTCGGTGGCTTGCATGCTGCTGGGTCTGTG 900
284 eRcysHisLeuIleProSerProThrPheAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCGGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTAA 950
301 AlaGlyLeuIleSerIleGlyGlyAlaLysCysLeuProValLysCysAs 317
951 CCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTACAACTTCACCT 1000
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGGTCTGCTGGAGAGATCATCTACATTTGCTGCTGGTCTGAT 1050
334 euLeuGlyLeuLeuGlyGlnIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTGGAGCGCGCAATGCGATGATTTGGCTTCCAGGTCCTCTAGCAT 1100
351 ThrValIlePheAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGGAATCAGCTTGGCATGCTGATAGCTCTCAGTCTGCTGCTCTGA 1150
367 eGlyGlyLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuT 384
1151 CAGGTTGCTCTTAATCTTAAATATGGAAGACGCTATGAGGCTAAA 1200
384 nGlyLeuLeuLeuAsnLeuLysIleThrPlyAlaProHisValAlaLys 400
1201 TATTTTGATGACAGTTTCTGGAAGTTTCTCATTTGGCTGGTGGAT 1250
401 TyrPheAspAspGlnValPheThrPlySerPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417

```

```

seq_documentation_block:
ID Q9UEC2 PRELIMINARY; PRT; 417 AA.
AC Q9UEC2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE RH BLOOD GROUP CE ANTIGEN POLYPEPTIDE.
GN RHCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kashwaza K., Ishikawa Y., Uchikawa M.;
RT "The serological profile and molecular basis of the RHE variants
(RHEFM and RHEKH) in Japanese.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018645; BAA33928.1; -.
DR InterPro; IPR001905; -.
DR InterPro; IPR002229; -.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRD.
SQ SEQUENCE 417 AA; 45505 MW; 67B53E74CB053DE CRC64;

```

```

alignment_scores:
Quality: 1928.00 Length: 417
Ratio: 4.869 Gaps: 0
Percent Similarity: 94.964 Percent Identity: 91.367

```

```

alignment_block:
US-09-600-714-41 x Q9UEC2 ..

```

```

Align seg 1/1 to: Q9UEC2 from: 1 to: 417

```

```

1 ATGAGCTTAAGTACCCGCGGTGTGCTGGCGGCTGCGCCCTGTGGGC 50
1 MetSerLysTyrProArgSerValArgArgCysLeuProLeuThrAl 17
51 CCTAACCTGGAGCAGCTTCATTTCTCTCTCTATTTTTPACCCACT 100
17 aLeuThrLeuGlnAlaIaLeuIleLeuLeuPheTyrPhePheThrHis 34
101 ATGAGCTTCTTACAGATGCAAAAGGGCTGCGTGCATCCATCAAGTT 150
34 yAspAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnVal 50
151 GGCCAAGATGACCGTATGCGGCGCATTTGGCTTGGGCTTCTCACCTC 200
51 GlyGlnAspLeuThrValMetAlaIaLeuGlyLeuGlyPheLeuThrSe 67
201 CAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGC 250
67 rAsnPheArgArgHisSerTyrPheSerValAlaPheAsnLeuPheMetL 84
251 TGGCGCTTGGTGTGAGTGGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
84 euAlaLeuGlyValGlnThrAlaIleLeuLeuAspGlyPheLeuSerGln 100
101 PheProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaTh 117
351 CATGAGTCTTTTGTGCTGCTGATCTCAGTGGATGCTGTGTTGGGGAAG 400
117 lMetSerAlaMetSerValLeuIleSerAlaGlyAlaValLeuGlyLysV 134
401 TCAACTGGCGCAGTTGGTGTGATGCTGCTGCTGAGAGTACAGCTTTA 450
134 aLaAsnLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeu 150

```

```
451 GGCACCTGAGGATGTCATCATTAATCTTCAACAGACATACACAT 500
|||||
151 GlyThrLeuThrMetValIleSerAsnIlePheAsnThrAspTyrHisMe 167
501 GAACATGATGCATCTACGTGTTCCGACGCTATTGCGCTGTGTGG 550
|||||
167 LAsnLeuArgHisPheTyrValPheAlaAlaTyrPheGlyLeuThrValA 184
551 CCTGTGCTCCCAAGGCTCTACCCGAGGAGGAGGATTAAGATCAG 600
|||||
184 LAspPheLeuProLysProLeuProLysGlyThrGluAspAsnSpGln 200
601 ACAGCAACGATACCCAGTTGTCTGCCATGCTGGGCGCCCTCTTGTG 650
|||||
201 ArgAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTr 217
651 GATGTTCTGGCCAAAGTTTCAACTGCTGCTCTGAGAAAGTCCAATCGAA 700
|||||
217 pMetPheTrpProSerValAsnSerProLeuLeuArgSerProIleGlnA 234
701 GGAAGATGCGCGTGTCAACACCTACTAGTGTGAGCAGTACGCGTGGTG 750
|||||
234 rGlySerAsnAlaMetPheAsnThrTyrTyrAlaLeuAlaValSerValVal 250
751 ACAGCCATCTCAGGGTCATCCTTGGCTCACCCCAAGGAGATCAGCAA 800
|||||
251 ThrAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMe 267
801 GACTTATGTGCACAGTCGCGGTGTGGCAGAGGCGGTGGCTGGTACT 850
|||||
267 tThrTyrValHisSerAlaValLeuAlaGlyGlyValAlaValAlaGlyThrS 284
851 CGTGTACACCTGATCCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||||
284 e rCysHisLeuIleProSerProTrpLeuAlaMetValLeuGlyLeuVal 300
901 GCTGGGCTGATCTCCGTGGGGGAGCAAGTACTCCGCGGGGTGTAA 950
|||||
301 AlaGlyLeuIleSerIleGlyAlaLysCysLeuProValCysCysAs 317
951 CCGAGTGTGGGGATCCCCACAGCTCATGCGGTACACTTCAGCT 1000
|||||
317 nArgValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerL 334
1001 TGCTGGGCTGCTGTGAGAGATCATCTACATTTGCTGCTGCTGCTGAT 1050
|||||
334 e uLeuGlyLeuLeuGlyGluIleThrTyrIleValLeuLeuValLeuHis 350
1051 ACCGTCGAGCCGCGCAATGCGATGATTGGCTTCAGGTCTCTCAGCAT 1100
|||||
351 ThrValITrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerI 367
1101 TGGGGAACCTACGCTTGGCCATCGTGAATAGCTTCACGCTGTGCTCTGA 1150
|||||
367 eGlyGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuTr 384
1151 CAGGTTTGCTCTCTAAATCTTAAATATGAAAGACCATGAGGCTAAA 1200
|||||
384 h rGlyLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLys 400
1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1250
|||||
401 TyrPheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPh 417
1251 T 1251
417 e 417
```

OM of: US-09-600-714-41 to: SwissProt_39:* out_format : pfs

Date: Sep 11, 2001 8:53 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame_n2p.model -DEV=xlp  
-O=/cgn2.1/USPRO.spool/US09600714/runat.11092001.083825.22207/app-query.fasta.1.1332  
-DB=SwissProt_39 -OFMT=fasta -SUFFIX=isp -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRAN=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORMEXT=0 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09600714.ecgn1.1_34 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPY  
-WAIT -THREADS=1
```

Search information block:

```
Query: US-09-600-714-41  
Query length: 1254  
Database: SwissProt_39:*  
Database sequences: 93435  
Database length: 34255486  
Search time (sec): 30.140000
```

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
SwissProt_39:RHD_HUMAN + 2118.00 2703.61	1.4e-143	416	0.02161	homo sapiens (human).	
SwissProt_39:RHLA_PATR + 1970.00 2514.36	5.0e-133	416	0.28814	pan troglodytes (chm).	
SwissProt_39:RHLA_GONG + 1948.00 2486.23	1.8e-131	416	0.28827	gorilla gorilla gorilla	
SwissProt_39:RHLA_PATR + 1947.00 2484.95	2.2e-131	416	0.28813	pan troglodytes (chm)	
SwissProt_39:RHLE_HUMAN + 1929.00 2461.93	4.2e-130	416	0.28817	homo sapiens (human).	
SwissProt_39:RHLE_PATR + 1914.00 2442.75	4.9e-129	416	0.28812	pan troglodytes (chm)	
SwissProt_39:RHLC_GONGO + 1903.00 2428.69	3.0e-128	416	0.28426	gorilla gorilla gorilla	
SwissProt_39:RHLA_MACU + 1642.00 2094.94	1.2e-109	416	0.28849	macaca mulatta (rhesus)	
SwissProt_39:RHLA_MACA + 1634.00 2084.71	4.3e-109	416	0.28841	macaca fascicularis (c)	
SwissProt_39:RHLA_HYLI + 1527.00 1949.30	1.8e-101	353	0.28446	hylobates pileatus (p)	
SwissProt_39:RHAG_HUMAN + 628.50 799.12	1.8e-37	409	0.20294	homo sapiens (human).	
SwissProt_39:RCHC_HUMAN + 141.00 171.16	0.0098	699	0.23327	homo sapiens (human).	
SwissProt_39:RCHC_HUMAN + 138.50 170.72	0.0143	507	0.23327	synochocystis sp. (st)	
SwissProt_39:RHUE_HYLI + 136.50 165.90	0.0204	660	0.20692	eschierichia coli. fct	
SwissProt_39:RHUE_HYLI + 135.50 165.90	0.0272	2358	0.29719	schizosaccharomyces	
SwissProt_39:RHUE_HYLI + 133.50 163.45	0.0354	521	0.20374	apis mellifera ligust	
SwissProt_39:RHUE_HYLI + 130.50 143.90	0.0644	3511	0.29424	mus musculus (mouse)	
SwissProt_39:RHUE_HYLI + 129.00 154.12	0.0715	852	0.16230	oryctolagus cuniculus	
SwissProt_39:RHUE_HYLI + 128.00 157.39	0.0800	501	0.15414	arabidopsis thaliana	
SwissProt_39:RHUE_HYLI + 123.50 137.16	0.1977	2715	0.29426	homo sapiens (human)	
SwissProt_39:RHUE_HYLI + 123.00 144.03	0.1968	1130	0.29747	mus musculus (mouse)	
SwissProt_39:RHUE_HYLI + 120.50 149.13	0.2695	429	0.25216	eschierichia coli. ars	
SwissProt_39:RHUE_HYLI + 120.00 148.10	0.2938	449	0.25192	mus musculus (mouse)	
SwissProt_39:RHUE_HYLI + 120.00 148.06	0.2940	451	0.25267	rattus norvegicus (rat)	
SwissProt_39:RHUE_HYLI + 117.50 145.31	0.4406	428	0.29290	eschierichia coli. pnc	
SwissProt_39:RHUE_HYLI + 116.50 142.86	0.5263	491	0.29290	ricettisia prowazeki	
SwissProt_39:RHUE_HYLI + 116.00 142.47	0.5696	477	0.29290	mycobacterium tuberculosis	
SwissProt_39:RHUE_HYLI + 115.50 133.33	0.6900	1446	0.29290	pseudorabies virus (p)	
SwissProt_39:RHUE_HYLI + 115.00 140.80	0.6741	499	0.29290	saccharomyces cerevisiae	
SwissProt_39:RHUE_HYLI + 115.00 139.69	0.6828	568	0.29290	rattus norvegicus (rat)	
SwissProt_39:RHUE_HYLI + 115.00 138.25	0.6943	672	0.29290	eschierichia coli. hyc	
SwissProt_39:RHUE_HYLI + 115.00 124.04	0.8181	3530	0.29290	homo sapiens (human)	
SwissProt_39:RHUE_HYLI + 114.50 141.99	0.7164	403	0.29290	eschierichia coli. hyc	
SwissProt_39:RHUE_HYLI + 114.00 141.20	0.7230	442	0.29290	synochocystis sp. (st)	
SwissProt_39:RHUE_HYLI + 114.00 141.40	0.7273	401	0.29290	alcaligenes eutrophus	
SwissProt_39:RHUE_HYLI + 113.50 139.81	0.8530	448	0.29290	homo sapiens (human)	
SwissProt_39:RHUE_HYLI + 113.50 139.02	0.8608	491	0.29290	metridium senile (brc)	
SwissProt_39:RHUE_HYLI + 113.50 139.00	0.8609	492	0.29290	saccharomyces cerevisiae	
SwissProt_39:RHUE_HYLI + 113.00 139.74	0.9198	419	0.29290	haemophilus influenzae	
SwissProt_39:RHUE_HYLI + 113.00 139.11	0.9265	451	0.29290	zymomonas mobilis. g	

SwissProt_39:ARSB_YEREN + 112.50 138.90	1.00	429	P74985	yersinia enterocoli
SwissProt_39:YDEE_ECOLI + 112.00 138.97	1.08	395	P31126	eschierichia coli.
SwissProt_39:GNTT_ECOLI + 111.50 137.46	1.18	437	P39835	eschierichia coli.
SwissProt_39:YCDG_ECOLI + 111.50 137.37	1.18	442	P75892	eschierichia coli.
SwissProt_39:HRG_HUMAN + 111.50 135.89	1.20	525	P04196	homo sapiens (human)

seq_name: SwissProt_39:RHD_HUMAN

```
seq_documentation_block:
ID RHD_HUMAN STANDARD: PRT; 416 AA.
AC 002161: 002162: 007618: 016355: 016147: 016235:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BLOOD GROUP RH(D) POLYPEPTIDE (RHESUS D ANTIGEN) (RHXXIII) (RH
DE POLYPEPTIDE 2) (RHPII).
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP TISSUE=Bone marrow;
RX MEDLINE=9306356; PubMed=1438298;
RA le van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
RA Cartron J.-P., Colin Y.;
RT "Molecular cloning and primary structure of the human blood group Rhd
RT polypeptide."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
RN 12
RP TISSUE=Bone marrow;
RX MEDLINE=9326085; PubMed=1379850;
RA le van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M.,
RA Cartron J.-P., Colin Y.;
RT "Multiple Rhd messenger RNA isoforms are produced by alternative
RT splicing."
RL Blood 80:1074-1078(1992).
RN 13
RP TISSUE=FROM N.A.
RX MEDLINE=9320449; PubMed=8329718;
RA Arce A.M., Thompson E.S., Wagner S., Coyne K.E., Ferdman B.A.,
RA Lublin D.M.;
RT "Molecular cloning of Rhd cDNA derived from a gene present in Rhd-
RT positive, but not Rhd-negative individuals."
RL Blood 82:651-655(1993).
RN 14
RP TISSUE=FROM N.A.
RX MEDLINE=93216282; PubMed=7916743;
RA Kaji E., Umenishi F., Iwamoto S., Ikemoto S.;
RT "Isolation of a new cDNA clone encoding an Rhd polypeptide associated
RT with the Rh blood group system."
RL Hum. Genet. 91:157-162(1993).
RN 15
RP TISSUE=FROM N.A.
RX MEDLINE=95329738; PubMed=7606008;
RA Huang C.H., Reid M.E., Chen Y.;
RT "Identification of a partial internal deletion in the RH locus
RT causing the human erythrocyte D-phenotype."
RL Blood 86:784-790(1995).
RN 16
RP TISSUE=FROM N.A. (SHORT FORM 1).
RX MEDLINE=9423583; PubMed=8180407;
RA Westhoff C.M., Wylie D.E.;
RT "Identification of a new Rhd-specific mRNA from K562 cells."
RL Blood 83:3098-3100(1994).
RN 17
RP TISSUE=FROM N.A. (SHORT FORM 2).
RX MEDLINE=94362249; PubMed=8080999;
RA Suyama K., Lunn R., Haller S., Goldstein J.;
RT "Rh(D) antigen expression and isolation of a new Rh(D) cDNA isoform
RT in human erythroleukemic K562 cells.";
```

RL	Blood 84:1975-1981(1994).
RN	[8]
RP	VARIANT BLOOD GROUP TAR.
RX	MEDLINE=95259709; PubMed=7741145;
RA	Rouillac C., le van Kim C., Boelet M., Carton J.-P., Colin Y.;
RT	"leu10pro substitution in the Rhd polypeptide is responsible for the Dvii category blood group phenotype.";
RL	Am. J. Hematol. 49:87-88(1995).
CC	-1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC	-1- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-1- TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING ERYTHROID CHARACTERS.
CC	-1- POLYMORPHISM: RHD AND RHCE ARE RESPONSIBLE FOR THE RH BLOOD GROUP SYSTEM. THE MOLECULAR BASIS OF THE TAR-RH40 BLOOD GROUP ANTIGEN IS A POLYMORPHISM IN POSITION 109.
CC	-1- SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHCE.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-slb.ch/announce/or_send_an_email_to_license@isb-slb.ch).
CC	-----
DR	EMBL; X63097; CAA44811.1; -
DR	EMBL; X63094; CAA44808.1; -
DR	EMBL; L08429; AAA02679.1; -
DR	EMBL; S57921; AAB26081.1; -
DR	EMBL; S78509; AAB34852.1; -
DR	EMBL; S70174; AAB30756.1; -
DR	EMBL; S73913; AAB31911.1; -
DR	PIR; S26564; S26564.
DR	MIM; 111680; -
DR	InterPro: IPR001905; -
DR	InterPro: IPR002229; -
DR	Pfam; PF00909; Anion_chan.transp. 1.
DR	PRINTS; PR000342; RHBST05SHD.
KM	Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing; Polymorphism.
FT	INIT MET 0 0
FT	TRANSMEM 11 31 POTENTIAL.
FT	TRANSMEM 43 63 POTENTIAL.
FT	TRANSMEM 76 96 POTENTIAL.
FT	TRANSMEM 106 126 POTENTIAL.
FT	TRANSMEM 129 149 POTENTIAL.
FT	TRANSMEM 166 186 POTENTIAL.
FT	TRANSMEM 202 222 POTENTIAL.
FT	TRANSMEM 237 257 POTENTIAL.
FT	TRANSMEM 286 306 POTENTIAL.
FT	TRANSMEM 333 353 POTENTIAL.
FT	TRANSMEM 377 377 POTENTIAL.
FT	VARSPPLIC 313 408 MISSING (IN SHORT ISOFORM 1).
FT	VARSPPLIC 315 315 C->S (IN SHORT ISOFORM 2).
FT	VARSPPLIC 316 416 MISSING (IN SHORT ISOFORM 2).
FT	VARIANT 109 L->P (IN TAR ANTIGEN).
FT	/FTID=VAR_006919.
FT	M->I.
FT	FTID=VAR_006920.
FT	CONFLICT 15 15 W->C (IN REF. 5).
FT	CONFLICT 38 38 E->G (IN REF. 4).
FT	CONFLICT 102 102 S->P (IN REF. 4).
FT	CONFLICT 126 126 V->A (IN REF. 4).
FT	CONFLICT 173 173 V->M (IN REF. 5).
FT	CONFLICT 181 181 G->T (IN REF. 4).
FT	CONFLICT 313 313 S->V (IN REF. 4 AND 7).
FT	CONFLICT 322 322 P->H (IN REF. 4).
FT	CONFLICT 397 397 E->V (IN REF. 5).
SO	SEQUENCE 416 AA; 45049 MW; 9136DF1A37D76B1B CRC64;

alignment_scores:	Quality:	Ratio:	Length:
Percent Similarity:	100.000	5.091	Gaps: 0
Percent Identity:	100.000	Percent Identity:	100.000
<p>alignment_block:</p> <p>US-09-600-714-41 x RHD_HUMAN ...</p>			
Align seg 1/1	to: RHD_HUMAN	from: 1	to: 416
4	ACGCTAAGTACCCCGGCTGTCTCCGGCGGCGCTGGCCCTTGCGGCT	53	
1	SerSerLysTyrProArgSerValAlaArgLysLeuProLeuTrpAlaLe	17	
54	AACACTGGAGCAGCTCTCATCTCTCTCTTATTTTATACCACTATG	103	
17	uThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrHisTyr	34	
104	ACGCTTCCTTAGAGATCAAAAGGGGCGCGGGATCCCTATCAAGTTGG	153	
34	spAlaSerLeuLyspGlnpLysGlyLeuAlaIleSerTyrGlnValGly	50	
154	CAGATCTGACCGGTGATGGCGGCAATGGCTTGAGGCTTCACCTTGAG	203	
51	GlnAspLeuThrValMetAlaAlaIleGlyLeuGlyPheLeuThrSer	67	
204	TTTCCGGAGACACACCTGGAGCAGTGTGGCTTAACTCTTCATGCTGG	253	
67	rPheArgTyrHisSerTrpSerSerValAlaPheAsnLeuPheMetLeu	84	
254	CGCTTGATGTGAGTGGGCAATCCGCTGGAGGCGCTTCAGACAGCTTC	303	
84	IaLeuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe	100	
304	CCCTTGGGAAAGGTGATCACACTGTCAGTATTCGGCTGGCCACCAT	353	
101	ProSerGlyLysValIleValIleThrLeuPheSerIleArgLeuAlaTrp	117	
354	GAGTGTCTTTGGGTCGTGATCTCAGTGGATGCGTCTTGGGAAAGTCA	403	
117	IserAlaLeuSerValIleLeuIleSerValAlaValLeuGlyLysVal	134	
404	ACCTGGCGCAGTTGGTGTGATGGTGGTGGGAGGAGACAGCTTATG	453	
134	snLeuAlaGlnLeuValIleMetValLeuValGluValThrAlaLeuGly	150	
454	AACCTGAGGATGTCATCATGTAATATCTTCAACAGACTACACATGAA	503	
151	AsnLeuArgMetValIleSerAsnIlePheAsnThrAspTyrHisMet	167	
504	CATGATGACATCTACGTTTCGGACGCTATTTTGGGCTGTCTGGCCT	553	
167	metMetHisIleTyrValIleAlaAlaTyrPheGlyLeuSerValAla	184	
554	GGTGGCTGGCAAGCCTTACCGGAGGAAAGAGATAAAGATCAGACA	603	
184	rPcysLeuProLysProLeuProGlnGlyThrGlnAspLysAspGlnThr	200	
604	GGAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGGAT	653	
201	AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTrp	217	
654	GTTTCGGCCAACTTTCACCTCTGCTCTCTCTAGAAAGTCATGCAAGA	703	
217	trPheTrpProSerPheAsnSerAlaLeuLeuArgSerTrpIleGlnArg	234	
704	AGATGCCCTGTTCAAACCTACTTGGCTGTAGACAGTACAGCGTGTACA	753	
234	ysAsnAlaValIleAsnThrTyrTyrAlaValAlaValSerValValThr	250	
754	GCACTTCAGGCTCATCTTGGCTCACCCCAAGGAAATCAGACAGAC	803	


```

251 AlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerLysTh 267
804 TTATGTGACAGTGGCGTGTGGAGAGCGCTGGCTGGTACCTGCT 853
|||||
267 rTyValHisSerAlaValIleAlaGlyGlyAlaValAlaGlyThrSec 284
884 GTACACCTATCCCTTCCTCCGCTGGCTGGCAATGGCTGGCTGGCT 903
|||||
284 ySHisLeuIleProSerProTrpLeuAlaMetValIleuGlyLeuValAla 300
904 GGGCGATCTCCGCTGGGGAGAGCAAGTACCTGGCGGGGTGTGAACCG 953
|||||
301 GlyLeuIleSerValGlyGlyAlaLysTyrLeuProGlyLysCysAsnAr 317
954 AGTGTGGGATTTCCCGACAGCTCCATCATGGGTACAACTTCACCTTGC 1003
|||||
317 gValLeuGlyIleProHisSerSerIleMetGlyTyrAsnPheserLeu 334
1004 TGGGTCTGCTTGGAGAGATCATCTACATTTGCTGCTGGCTTGAATAC 1053
|||||
334 euGlyLeuLeuGlyGlyIleIleTyrIleValLeuLeuValLeuAspThr 350
1054 GTGGGAGCGCGCAATGCATGTGGCTTCGAGCTCCCTCAGCATTCG 1103
|||||
351 ValGlyAlaGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerIleG 367
1104 GGAATCAGCTGGCCATCGTGTAGTACCTCTCAGCTGGCTGCTCAGAC 1153
|||||
367 yGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuThr 384
1154 GTTGTCTCTTAATCTTAATAATATGAAAGCACTCATGAGGCTTAATAT 1203
|||||
384 yLeuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLysTyr 400
1204 TTTGATGACCAAGTTTCTGGAAGTTCTCATTTGGCTGGTGAATTT 1251
|||||
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416
seq_name: SwissProt_39:RHLR_PANTR

seq_documentation_block:
ID RHLR_PANTR STANDARD; PRT; 416 AA.
AC Q28B14;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE RH-LIKE PROTEIN IIR (RHESUS-LIKE PROTEIN IIR).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95083595; PubMed=7993375;
RA Salignon I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: L37050; AAA65624.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
KW Erythrocyte; Transmembrane.
FT INIT_MET 0
FT TRANSMEM 11 31
FT TRANSMEM 43 63
FT TRANSMEM 76 96
FT TRANSMEM 124 144
FT TRANSMEM 171 191
FT TRANSMEM 202 222
FT TRANSMEM 237 257
FT TRANSMEM 264 284
FT TRANSMEM 286 306
FT TRANSMEM 330 350
FT TRANSMEM 357 377
SQ SEQUENCE 416 AA; 45299 MW; C055D7CC8BA0420 CRC64;

alignment_scores:
Quality: 1970.00 Length: 416
Ratio: 4.888 Gaps: 0
Percent Similarity: 96.875 Percent Identity: 93.029

alignment_block:
US-09-600-714-41 x RHLR_PANTR ..

Align seg 1/1 to: RHLR_PANTR from: 1 to: 416

4 ACCTTAAGTACCCGGGCTGTCCGGCGCTGCTGCCCTGGGACCT 53
|||||
1 SerSerLysTyrProArgSerValAlaArgGlySerProLeuGlyAlaIle 17
54 AACACTGGAAGCAGCTCATCTCTCTCTCTATTTTACCACATAG 103
|||||
17 uThrLeuGlnAlaAlaLeuIleLeuLeuPheTyrPhePheThrGlnTyrA 34
104 AGCTTTCTTAGAGATCAAAAGGGCTCGTGCGATCCATCAAGTTGC 153
|||||
34 sPAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnValGly 50
154 CAAGATCGACCGGTGATGGCGCAATGGCTGGGGTTCCTACCTCGAG 203
|||||
51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
204 TTTCGGAGACACAGCTGAGCAGTGTGGCTTCAACCTTCATGCTGG 253
|||||
67 rPheAlrArgHisSerTrpSerSerValAlaPheSerLeuPheMetLeuA 84
254 CGCTTGCTGTGAGTGGGCAATCCGCTGAGAGGCTTCTGAGCCAGTTC 303
|||||
84 lAlaLeuGlyAlaGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTGGGAAGGTGATCATCACATGTTCACTATTCGGCTGGCCACCAT 353
|||||
101 ProProGlyLysValValIleThrIleuPheSerIleArgGlyAlaIleThr 117
354 GAGTCTTTGTGGTGTGATCTCAAGTGAGTCTGTGGGGAAGGTCA 403
|||||
117 rSerAlaLeuSerValIleIleSerValAlaSprAlaValLeuGlyLysValA 134
404 ACTTGCGCAGTGTGGTGTGATGGTGGCTGGTGGAGTGCACCTTAGGC 453
|||||
134 sIleValGlnIleuValAlaMetValIleuValGlnValThrAlaLeuGly 150
454 AACCTGAGAGTGCATCATGATATCTTCAACAGACAGTACCATGAA 503
|||||
151 ThrValArgMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
504 CATGATGCACATCTACGTTGTGCAAGCTATTTTGGGCTGTGTGGCCT 553

```

```

167 nleuethisiletyValPhealaIatyrPheyleuSerValaIat 184
554 GGTGGCTGCCAAAGCCTTACCAGGAGGAGGATTAAGATCAGCA 603
184 rPcysleuProlyserProleupPolyglyThrGluAspLysAspGlnIle 200
604 GCAACGATACCAGTGTGTGTCGACGCTGGCGCCCTTCTTGAT 653
201 AlaThrIleProserleuSerAlaMetleuGlyAlaLeuPheleuTrpMe 217
654 GTTGGCCCAAGTTTCAACTGCTGCTGCTGAGAAAGTCCAAATCGAAGA 703
217 tPheTrpProSerPheAsnSerAlaLeuLeuArgSerProIleGluArgL 234
704 AGAATGCCGTGTTCACACCTACTAGCTGTAGCAGGAGGAGGAGGAGGAG 753
234 ySAsnAlaValAlaPheAsnThrTyrrAlaValAlaValaSerValThr 250
754 GGCATCTGAGGATCATCTGCTGCTACCCCAAGGAGGAGGAGGAGGAGGAG 803
251 AlaIleSerGlySerSerSerLeuAlaHisProGlnGlyLysIleSerMet 267
804 TTAATGACACAGTGGGCTGTGGCAGAGGCGCTGCTGGGCTGCTGCT 853
267 rTyMetHisAsnAlaValaLeuAlaGlyValaValaValaGlyThrSerC 284
854 GTACACCTATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
284 ySHisLeuIleProserProTrIleuAlaMetValaLeuGlyLeuValaAla 300
904 GGGCTGATCTCCGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 953
301 GlyLeuIleSerValaGlyValaLysTyrrLeuProGlyCysAsnAr 317
954 AGTGTGGGAGTATCCCAAGCTCATGATGGGCTGACAACTTCAAGCTTGC 1003
317 gValleuGlyIleProHisSerSerValaMetGlySerAsnPheSerTrpL 334
1004 TGGCTCTGCTGGAGAGATCATCTGATTTGCTGCTGCTGCTGCTGCTGAT 1053
334 euGlyLeuLeuGlyGluIleIleTyrrIleValleuValaValaArgHisThr 350
1054 GTCCGAGCCGCAATGCGATGATGGCTTCCAGCTCTCTCCAGCATTTGG 1103
351 IleTrpAsnGlyAsnGlyMetIleGlyPheGlnValleuLeuArgIleG 367
1104 GGAATCTAGCTGGCAGTCGATAGCTTCACGCTGCTGCTGCTGCTGAGAG 1153
367 yGluPheSerleuAlaThrTrIleAlaLeuThrSerGlyLeuLeuThrG 384
1154 GTTGGCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTTAATAT 1203
384 lYleuLeuLeuAsnLeuLysIleTrpLysAlaProHisGlnAlaLysTy 400
1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
401 PheAspAspArgIleValaPheTrpLysPheProHisLeuAlaValaGlyPhe 416
seq_name: SwissProt_39:RHLD_GORGO
seq_documentation_block:
ID RHLD_GORGO STANDARD; PRT; 416 AA.
AC 028427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE RHESUS-LIKE PROTEIN.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9595;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=95085595; Pubmed=7993375;
RA Salvagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE RH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC
CC DR EMBL: L37053; AA65627.1;
CC DR InterPro: IPR001905;
CC DR Pfam: PF00909; Ammonium_transp. 1.
CC DR PRINTS: PR00342; RHESUSRHD.
CC KM Erythrocyte; Transmembrane.
CC FT INIT MET 0 BY SIMILARITY.
CC FT TRANSMEM 11 31 POTENTIAL.
CC FT TRANSMEM 43 63 POTENTIAL.
CC FT TRANSMEM 76 96 POTENTIAL.
CC FT TRANSMEM 124 144 POTENTIAL.
CC FT TRANSMEM 171 191 POTENTIAL.
CC FT TRANSMEM 202 222 POTENTIAL.
CC FT TRANSMEM 237 257 POTENTIAL.
CC FT TRANSMEM 264 284 POTENTIAL.
CC FT TRANSMEM 286 306 POTENTIAL.
CC FT TRANSMEM 330 350 POTENTIAL.
CC FT TRANSMEM 357 377 POTENTIAL.
CC SQ SEQUENCE 416 AA; 44838 MW; 398877E2B7BEE802 CRC64;

```

```

alignment_scores:
Quality: 1948.00 Length: 416
Ratio: 4.870 Gaps: 0
Percent Similarity: 96.154 Percent Identity: 92.788

```

alignment_block:

US-09-600-714-41 x RHLD_GORGO ..

Align seg 1/7 to: RHLD_GORGO from: 1 to: 416

```

4 AGCTCTAGATACCCGGGCTGTGTCGGGCGCTGCTGCTGCTGCTGCT 53
|||||
1 SerSerLysTyrrProArgSerValaArgCysLysLeuProLeuCysAlaLe 17
54 AACACGTGAACAGCTGCTCATCTCTCTTCTATTTTTCACCATG 103
|||||
17 uThrLeuGlnAlaAlaLeuThrLeuLeuPheTyrrPhePheThrHisTy 34
104 ACAGCTTCCTTAGAGATCAAAAGGGGCTGTCGATCCATCAAGTTGGC 153
|||||
34 spAlaSerleuGluAspGlnLysGlyLeuValaAlaSerTyrrGlnValGly 50
154 CAACATCTGACCGTGATGCGCGCATTTGGCTTGGCTTCTCACCTGAG 203
|||||
51 GlnAspLeuThrValaMetAlaAlaIleGlyPheGlyPheLeuThrSer 67
204 TTTCCGAGACACAGCTGAGACAGTGGCGCTTCAACCTGTATGCTGG 253
|||||
67 rPheArgGlyHisSerTrpSerSerValaAlaPheAsnLeuPheMetLeuA 84

```

```

|||||
384 LYLEULEAAsnleuAysllEtRpySAIApRohISAlAAlAysIYR 400
1204 TTTCGATGACCAAGTTTCTGGAAAGTTTCCTCAATTTGGCTGTGGATT 1251
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416

seq_name: SwissProt_39:RHLA_PANTR

seq_documentation_block:
ID RHLA_PANTR STANDARD; PRT; 416 AA.
AC 028813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN IA (RHESUS-LIKE PROTEIN IA).
OS Pan troglodytes (Chimpusae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RC MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Collin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC
EMBL: L37049; AAA65623.1; -.
DR InterPro: IPR001905; -.
DR Pfam: PF00909; Ammonium transp. 1.
DR PRINTS: PR00342; RHESUSRHD.
KW Erythrocyte; Transmembrane.
KW INT_MET 0
FT FT 11 31 POTENTIAL. BY SIMILARITY.
FT FT 43 63 POTENTIAL.
FT FT 76 96 POTENTIAL.
FT FT 124 144 POTENTIAL.
FT FT 171 191 POTENTIAL.
FT FT 202 222 POTENTIAL.
FT FT 237 257 POTENTIAL.
FT FT 264 284 POTENTIAL.
FT FT 286 305 POTENTIAL.
FT FT 330 350 POTENTIAL.
FT FT 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 45461 MW; B0B566734DB5E14D CRC64;

alignment_scores:
Quality: 1947.00 Length: 416
Ratio: 4.867 Gaps: 0
Percent Similarity: 96.154 Percent Identity: 92.548

alignment_block:
US-09-600-714-41 x RHLA_PANTR ..
Align seg 1/1 to: RHLA_PANTR from: 1 to: 416

```

Length: 416
Gaps: 0
Percent Identity: 92.5488

Align seg 1/1 to: RHLA_PANTR from: 1 to: 416

```

4  AGCTTAAGTACCGCGGTGTGTCGCGGCGTCCGCCCTTGGGCGCT 53
   |||||||
1  SerSerIstYrProArgSerValArgArgCysLeuProLeuGlyAlaLe 17
54  AACACTGGAAGAGCTCATCTTCCTCTCTCATTTTATTTTACCCACTATG 103
   |||||||
17  uThrLeuGlnAlaAlaLeuLeuLeuPheThrPhePheThrGlnTyrA 34
104  ACGCTTCCCTTAGAGAGATCAAAAGGGCTGTCGATCCATCAAGTTGAC 153
   |||||||
34  sPAlaSerLeuGlnuSpGlnuSgLyLeuValAlaSerTyrGlnValGly 50
154  CAAGATCTGACCGTGAAGGGCGCATTTGGCTTCCCTCCACCTCGAG 203
   |||||||
51  GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerS 67
204  TTTCCGGGACACAGCTGAGAGAGTGTGCGCTTCAACCTCTTCATGCTGG 253
   |||||||
67  rPheArgIstHisSerTyrSerSerValAlaIaPheSerLeuPheMetLeuA 84
254  CGCTTGTGTGAGTGGCAATCTGCTGAGACGGCTCTCGTACAGCAGTTC 303
   |||||||
84  lAlaGlyValAlaGlnTrpAlaIleLeuLeuAspGlyPheLeuSerInPhe 100
304  CCTTGTGGGAAGTGGTCAATCAGACTGTTCAATTTGGCTGGCCAGCAT 353
   |||||||
101  ProProGlyLysValValIleThrLeuPheSerIleArgLeuAlaThrTh 117
354  GAGTGTCTTGGGTGGATGATCTGATGCTGCTCTTGGGGAAGTCA 403
   |||||||
117  rSerAlaLeuSerValIleLeuSerValAspAlaValAlaGlyLysValA 134
404  ACTTGGCCAGTGTGTGATGATGCTGCTGTCGAGAGTACAGCTTACAGC 453
   |||||||
134  snLeuValAlaGlnLeuValAlaMetValIleuValAlaGlnValThrAlaLeuGly 150
454  AACCTGAGGATGTCATCATGTAATATCTTCAACACAGACTACCAATGAA 503
   |||||||
151  ThrValArgMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
504  CATGATGACATCTACGCTTCCGACGCTATTTTGGCTGTCTGTGGCGCT 553
   |||||||
167  nLeuMetHisIleTyrValPheAlaAlaTyrPheGlyLeuSerValAlaIat 184
554  GGTGCTGCCAAAGCCTTACCCGAGGAGACGAGATTAAGATAGACA 603
   |||||||
184  rPcysLeuProLysProLeuProLysGlyThrGlnAspLysAspGlnIle 200
604  GCACGATACCCAGTTTGTGCTGCTGAGCGCGCTCTTCTTGTGAT 653
   |||||||
201  AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuThrPme 217
654  GTTTCGGCGCAAGTTTCAACTGCTGCTCTGAGAGTCCATTCGAAAGGA 703
   |||||||
217  tPheThrProSerPheAsnSerAlaLeuLeuAlaArgSerProIleGlnuAgl 234
704  AGAATGCGGTGTCAACACTACTAGTGTAGAGAGTACGAGGTGTGAGCA 753
   |||||||
234  ysAsnAlaValAlaPheAsnThrTyrTyrAlaValAlaValSerValValThr 250
754  GCCATCTCAGGGTCATCTTGGCTACCCCAAGGAAGATCAGCAAGAC 803
   |||||||
251  AlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerMetSe 267
804  TTAATGTGACAGTCCGCTGTGTCGAGAGAGCGGTGTGGTACCTGCTGT 853
   |||||||
267  rTyrMetHisAsnAlaValAlaLeuAlaGlyGlyValAlaValAlaGlyThrSerC 284
854  GTCACTGATCCCTTTCCTGCGTGGCTGCGAGGAGTGGCTGGCTGTGGCT 903
   |||||||
284  ysHisIleuIleThrSerProThrPheuAlaMetValLeuGlyLeuValAla 300

```

```

904  GGGCTGATCTCCGTCGGGGGAGCCAGTACCTGCCGGGGTGTGAACCG 953
   |||||||
301  GlyLeuIleSerIleGlyAlaAlaLysTyrLeuProGlyCysCysAsnAr 317
954  AGTGGTGGGANTCCCCACAGCTCATCATGAGGCTACAACTTCAAGCTTGC 1003
   |||||||
317  gValLeuGlyIleTyrHisSerSerValMetHisTyrAsnPheSerLeuL 334
1004  TGGGCTGCTTGGAGAGATCATCTACTATTTGCTGTGCTGCTGTGATPAC 1053
   |||||||
334  euGlyLeuLeuGlyGlyIleIleTyrIleValAlaLeuValAlaHisHisThr 350
1054  GTCGGAGCCGGCAATGGCATGATTTGGCTTCCAGGCTCTCTACACATTGG 1103
   |||||||
351  ValTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuArgIleG 367
1104  GGAATCTGAGCTTGGCCATCGATAGCTCTTCACGCTGTGCTCTGACAG 1153
   |||||||
367  yGluPheSerLeuAlaThrThrIleAlaLeuThrSerTyrLeuLeuThrG 384
1154  GTTGTCTCTTAATCTTAAATATGAAAGACCTCATGAGCTAAATAT 1203
   |||||||
384  lYleuLeuLeuAsnLeuLysIleTyrPylSalapronHisAlaAlaLysTyr 400
1204  TTTGATGACCAAGTTTCTTGGAAGTTTCTCATTTGGCTGTGATTT 1251
   |||||||
401  PheAspAspGlnValPheThrPylsPheProHisLeuAlaValGluPhe 416
seq_name: SwissProt_39: RHCE_HUMAN
seq_documentation_block:
ID      RHCE_HUMAN          STANDARD:      PRT:   416 AA.
AC      P18577; Q02163; Q02164; Q02165; Q16160;
DT      01-NOV-1990 (Rel. 16, last sequence update)
DT      01-NOV-1990 (Rel. 16, last sequence update)
DT      01-OCT-2000 (Rel. 40, last annotation update)
DE      BLOOD GROUP RH(CE) POLYPEPTIDE (RHESUS C/E ANTIGENS) (RH30A) (RH1XB)
DE      (RH POLYPEPTIDE 1) (RHP1).
GN      RHCE OR RHC OR RHE.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=Bone marrow;
RX      MEDLINE=90349591; PubMed=1696722;
RA      Cherif-Zahar B., Bloy C., le van Kim C., Blanchard D., Bailly P.,
RT      Hermand P., Salmon C., Carton J.-P., Colin Y.;
RT      "Molecular cloning and protein structure of a human blood group Rh
RT      polypeptide.";
RL      Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91058522; PubMed=2123099;
RA      Avent N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;
RT      "cDNA cloning of a 30 kDa erythrocyte membrane protein associated
RT      with Rh (Rhesus)-blood-group-antigen expression.";
RL      Biochem. J. 271:821-825(1990).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93216282; PubMed=7916743;
RA      Kajii E., Umenishi F., Iwamoto S., Ikemoto S.;
RT      "Isolation of a new cDNA clone encoding an Rh polypeptide associated
RT      with the Rh blood group system.";
RL      Hum. Genet. 91:157-162(1993).
RN      [4]
RP      SEQUENCE FROM N.A. (FORMS RHIV, RHVI AND RHVII).
RC      TISSUE=Bone marrow;
RX      MEDLINE=92360855; PubMed=1379850;
RA      le van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M.,
RA      Carton J.-P., Colin Y.;
RT      "Multiple Rh messenger RNA isoforms are produced by alternative

```



```

204 TTTCGGAGACACAGCTGGAGACAGTGTGGCCTTCAACCTTTCATGCTGG 253
    :|||||
67 nPheArGhrHisSerTrpSerSerValAlaPheSnlLeuPheMetLeuA 84
254 CGCTTGGTGTGACAGTGGCAATCCTGCTGGACGGCTTCTTCAAGCAGTTC 303
    :|||||
84 lAlenGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTGTGGAGAGTGGTGCATCACACGTGTTCAGTATTGGCTGGCCACCAT 353
    :|||||
101 ProProGlyLysValValIleThrLeuPheSerIleArgGlyLeuAlaThrMe 117
354 GAGTGTCTTTCGCTGCTGATCTCAGTGAGATCTGTCTTGGGGAGAGTCA 403
    :|||||
117 tSerAlaMetSerValIleLeuIleSerAlaGlyAlaValLeuGlyLysValA 134
404 ACTTGGCCCACTGTGGTGTGATGGTGGTGGAGGTGACAGCTTTAGGC 453
    :|||||
134 snLeuAlaGlnLeuValValMetValLeuValGluValThrAlaLeuGly 150
454 AACCTGAGAGTGCATCAGTAATATCTTCAACAGAGCTACCCAGTAA 503
    :|||||
151 ThrLeuArgMetValIleSerAsnIlePheAsnThrAspTrpHisMetAs 167
504 CATGATGCACATCTACGTGTTCGAGCCTATTGGGCTGTGTGGCCT 553
    :|||||
167 nLeuArgHisPheTrpValPheAlaAlaTrpPheGlyLeuThrValAlaTr 184
554 GGTGCTGCCCAAGCCTCTACCCGAGGAGACGAGAGTAAGATCAGACA 603
    :|||||
184 rPcysLeuProLysProLeuProLysGlyThrGlnAspAspAlaPdnArg 200
604 GCAACGATACCAGATTGTCTGCCATGCTGGGGCGCCCTCTCTGTGGAT 653
    :|||||
201 AlAlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTrpMe 217
654 GTTCTGGCCAAAGTTTCACTGTCTGCTGTGAGAGTCCAAATCGAAGA 703
    :|||||
217 tPheTrpProSerValAsnSerProLeuLeuArgSerProIleGlnArgL 234
704 AGAATGCCGCTTTCACACCTACTATGTGTAGACAGTACGGTGTGTGCA 753
    :|||||
234 ySaSnAlaMetPheAsnThrTrpAlaLeuAlaValSerValValThr 250
754 GGCATCTCAGAGGTCATCCTTGGCTACCCCAAGGAGATCAGACAAGAC 803
    :|||||
251 AlAlaIleSerGlySerSerLeuAlaHisProGlnArgLysIleSerMetTh 267
804 TTATGTGCACAGTGCAGTGTGGAGAGGCGTGTGGGTACCTGCT 853
    :|||||
267 rTyValHisSerAlaValLeuAlaGlyGlyValAlaValAlaIlyThrSerC 284
854 GTCAACCTGATCCCTTCCGCTGGCTGGCATGCTGGTGGCTGTGGCT 903
    :|||||
284 yShSLeuIleIleProSerProTrpLeuAlaMetValLeuGlyLeuValAla 300
904 GGGCTGATCTCCGTCGGGGAGCAAGTACTCGCGGGGTGTGTAAACCG 953
    :|||||
301 GlyLeuIleSerIleGlyGlyAlaLysCysLeuProValCysCysAsnArg 317
954 AGTCGCTGGGATTCGCCACAGCTCCATCATGGGCTACAACTTCAGCTTGC 1003
    :|||||
317 gValLeuGlyIleHisHisIleSerValMetHisSerIlePheSerLeuL 334
1004 TGGGTCTGCTGGAGAGATCATCTACATGTGTGCTGCTGCTGATACC 1053
    :|||||
334 euGlyLeuLeuGlyGlyIleThrTrpTrpIleValLeuLeuValLeuHisThr 350
1054 GTCCGAGCGCGCAATGGCATGATTGGCTTCAGGCTCCCTCAGACATGG 1103
    :|||||
351 ValTrpAsnGlyAsnGlyMetIleGlyPheGlnValLeuLeuLeuSerIleG 367

```

```

1104 GGAACGACGCTTGGCCATCGTATAGCTCTCAGCTGTGCTCTCCAGAG 1153
    :|||||
367 yGluLeuSerLeuAlaIleValIleAlaLeuThrSerGlyLeuLeuThr 384
1154 GTTGTCTCTTAATCTTAATATGGAAGCAGCTCATGAGCTTAATAT 1203
    :|||||
384 lYleuLeuLeuAsnLeuLysIleTrpLysAlaProHisValAlaLysTyr 400
1204 TTTGATGACCAAGTTTCTGGAAGTTTCCTCATTTGGCTGTGGATTT 1251
    :|||||
401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416

```

seq_name: SwissProt_39:RHLE_PANTR

seq_documentation_block:

```

ID   RHLE_PANTR          STANDARD;              PRT;          416 AA.
AC   O28812;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   RH-LIKE PROTEIN IIF (RHESUS-LIKE PROTEIN IIF).
OC   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9505595; PubMed=7993375;
RA   Salvignol Y., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA   Collin Y., Ruffie J.;
RT   "Molecular genetics of chimpanzee Rh-related genes: their
RT   relationship with the R-C-E-F blood group system, the chimpanzee
RT   counterpart of the human rhesus system.";
RL   Biochem. Genet. 32:201-221(1994).
CC   - FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC   HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC   - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   - SIMILARITY: BELONGS TO THE RH FAMILY.
CC   CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL: L37048; AAA5622.1; -
DR   InterPro: IPR001905; -
DR   InterPro: IPR002229; -
DR   Pfam: PF00909; Ammonium_transp. 1.
DR   PRINTS: PR00342; RHESUSRHD.
KW   Erythrocyte; Transmembrane.
FT   INIT_MET              0
FT   TRANSMEM              11
FT   TRANSMEM              43
FT   TRANSMEM              76
FT   TRANSMEM              124
FT   TRANSMEM              171
FT   TRANSMEM              202
FT   TRANSMEM              237
FT   TRANSMEM              264
FT   TRANSMEM              286
FT   TRANSMEM              330
FT   TRANSMEM              357
SQ   SEQUENCE              416 AA; 45136 MW; 325ED916268BAF2F CRC64;

```

alignment_scores:

```

Quality: 1914.00      Length: 416
Ratio: 4.833          Gaps: 0
Percent Similarity: 95.192      Percent Identity: 91.106

```

alignment_block:
US-09-600-714-41 x RHLE_PANTR ..

Align seg 1/1 to: RHLE_PANTR from: 1 to: 416

```

4 AGCTTAAGTACCGCGGTCTGTCCGGCGCTCCGCCCTCTGGCCCT 53
  |||||
1 SerSerLysTyrProArgSerValArgCysLeuProLeuGlyAlaLe 17
54 AACACTGAAGAGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 103
  |||||
17 ThrLeuGlnAlaAlaLeuLeuLeuPheThrPhePheThrHisTyrA 34
104 AGCTTCTTCTAGAGGATCAAAAGGGCTCTGTCATCTCTCTCTG 153
  |||||
34 SpAlaSerLeuGlnLysPglInySgLYleuValAlaSerTyrGlnValGly 50
154 CAAGATCTGACCGTGAATGGCGCATTTGGCTTGGCTTCTCAGCT 203
  |||||
51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
204 TTTCGGAGACACAGCTGAGCAGTGTGGCTTCAACCTCTCATGCTG 253
  |||||
67 PheArgAlaHisSerTyrPheSerSerValAlaPheAsnLeuPheMetLeuA 84
254 CGCTTGTGTGCAAGTGGCAATCTCTGTGAGCGCTTCTCTGAGCCAGTTC 303
  |||||
84 LeuGlnLYalGlnThrAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTTTGGGAAGGTGTCTACATCAGTTCATTTCCGCTGGCCACCAT 353
  |||||
101 PropGlyLysValValIleThrLeuPheSerIleArgLeuAlaThrH 117
354 GAGTCTTGTGGTGTGATCTCATGATGATGCTGTCTTGGGGAAGGTA 403
  |||||
117 rSerAlaLeuSerValLeuIleSerAlaGlyAlaValLeuGlyTyrValA 134
404 ACTTGCGCAGTTGGTGTGATGATGATGATGATGATGATGATGATG 453
  |||||
134 snLeuValGlnLeuValValMetValLeuValGlnValThrAlaLeuGly 150
454 AACCTGAGAGATGTCATCATATATCTTCAACACAGATACACATGAA 503
  |||||
151 ThrMetArgMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
504 CAGTATGACATCTACGTGTGACGACCTATTGGTGGCTGTGTCGCTG 553
  |||||
167 MetLeuThrHisPheTyrLeuPheThrAlaTyrPheGlyValIleThrValAla 184
554 GGTGCTGCAAGAGCTCTACCCGAGGGAAGGAGATAAAGATCAAGACA 603
  |||||
184 rPcysLeuProLysProLeuProAspValLYsgLnsPlysAspGlnIle 200
604 GCAAGATATCCAGATTGTCTGCCATCTGGGCGCCCTTCTTGTGAT 653
  |||||
201 AlaThrIleProSerLeuSerAlaMetLeuGlyAlaLeuPheLeuTyrPhe 217
654 GTTTCGGCAAGTTCACTGCTGTGCTGTGAGAACTCAATCGAAGA 703
  |||||
217 tPheTyrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnArgL 234
704 AGAATCCCTGTTCACACCTACTATGCTGTAGACATCAGCGTGTGACA 753
  |||||
234 yAsnAlaValAlaPheAsnThrTyrTyrAlaLeuAlaValSerValValThr 250
754 GCCATCTCAGGGTCATCTTGGCTCACCCCAAGGAGATCAGACAGAC 803
  |||||
251 AlaIleSerGlySerSerLeuAlaHisProGlnGlyLysIleSerMetH 267
804 TTATGTGACAGTGGGTGTGGCAGAGCGGCTGTGGTGAACCTCGT 853
  |||||
267 rTyrValHisSerAlaValLeuAlaGlyLYalValAlaValGlyThrSerC 284

```

```

854 GTCACCTGATCCCTTCTCCGGCTTGCACATGCTGCTGGGCTTGTGCT 903
  |||||
284 yHisIleuIleProSerProThrLeuAlaMetValLeuGlyLeuValAla 300
904 GGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCCGGGGTGTGTACCG 953
  |||||
301 GlyLeuIleSerIleGlyLYalAlaLysCysGlyProGlyCysCysAsnAr 317
954 AGTGTGGGATTTCCCAAGCTCATGAGGATGAGGATCACTGACCTGAC 1003
  |||||
317 gValLeuGlyIleProAspSerSerValMetHisTyrAsnPheSerLeuL 334
1004 TGGTCTGCTTGGAGAGATCATCTACATTTGCTGCTGCTGCTGCTG 1053
  |||||
334 euGlyLeuLeuGlyLysIleIleTyrIleValLeuValAlaArgHisThr 350
1054 GTGGAAGCCGCGCAATGAGCATGATTTGGCTTCCAGGCTTCTCAGAT 1103
  |||||
351 ValTyrPasnGlyAsnGlyMetIleGlyPheGlnValLeuLeuSerMetL 367
1104 GGAACCTCAGCTTGGCCATGCTGATAGCTCTCAGCTGCTGCTGCTG 1153
  |||||
367 yGluLeuSerLeuAlaIleAlaIleAlaLeuThrSerGlyLeuLeuThg 384
1154 GTTTCCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTAAATAT 1203
  |||||
384 LYleuLeuLeuAsnLeuLysIleTyrPlyAlaProHisValAlaLysTyr 400
1204 TTGTATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGATTT 1251
  |||||
401 SerAspAspGlnValAlaPheThrPlysPheProHisLeuAlaValGlyPhe 416

seq_name: SwissProt_39:RHLC_GORGO
seq_documentation_block:
ID RHLC_GORGO STANDARD: PRT: 416 AA.
AC Q28426;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-LIKE PROTEIN IC (RHESUS-LIKE PROTEIN IC).
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Collin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human Rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
RL -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L37052; AAA65626.1; -
DR InterPro: IPR001905; -
DR InterPro: IPR002229; -
DR Pfam: PF00909; AminoLum_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.

```

KM Erythrocyte; Transmembrane.
 FT INT. MET 0 0 BY SIMILARITY.
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 43 63 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 124 144 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 237 257 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 SQ SEQUENCE 416 AA; 45129 MW; 0F3B623F908E087 CRC64;

alignment_scores:
 Quality: 1903.00 Length: 416
 Ratio: 4.806 Gaps: 0
 Percent Similarity: 95.192 Percent Identity: 89.904

alignment_block:
 US-09-600-714-41 x RHLG-GORGO

Align seg 1/1 to: RHLG-GORGO from: 1 to: 416

4 AGCTTAAGTACCGCGGTCTGTCGGCGGCTGCGCCCTGCGCCCT 53
 1 SerSerLysTyrProArgSerValArgCysLeuProLeuCysAlaIle 17
 54 AACACTGGAAGCAAGCTTCATTCTCTCTATTTTATTTTATCCCACTATG 103
 17 uThrLeuGlnAlaLeuIleLeuLeuPheThyRphSerThrHisTyrA 34
 104 ACGCTCTCTAGAGATCAAAAGGGCTCGTGGCATCTCATCAATGGC 153
 34 spAlaSerLeuGlnAspGlnLysGlyLeuValAlaSerTyrGlnValGly 50
 154 CAAGATCGACCGTATGGCGCATGGCTGGCTGCGCTCCACCTCGAG 203
 51 GlnAspLeuThrValMetAlaAlaIleGlyPheGlyPheLeuThrSerSe 67
 204 TTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAACCTTTTCATGCTGG 253
 67 rPheArgAlaHisSerTyrSerSerValAlaPheAsnLeuPheMetLeuA 84
 254 GCGTTGGTGTGAGTGGCAATCTGCTGAGAGCTTCTGAGCCAGTTC 303
 84 lAlaGlyValGlnTrrPalaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
 304 CTTTCTGGGAAGTGGTGCATCACACTGTTCAGTATTCGGTGGCCACAT 353
 101 ProProGlyLysValValIleThrLeuPheSerIleArgGlnAlaIleThr 117
 354 GAGTCTTTTGTGCTGTGATCTCAGTGAATGCTGTCTTGGGGAAGTCA 403
 117 tSerAlaLeuSerValIleLeuIleSerAlaGlyAlaValLeuGlyTyrValA 134
 404 ACTTGGCCAGTGTGGTGGATGGTGGTGGAGTGGAGTGGAGTTCATGCG 453
 134 snLeuValGlnLeuValAlaMetValGlnValAlaThrAlaLeuGly 150
 454 AACCTGAGATGATCATCATATATCTTCAACAGACAGTACCATGAA 503
 151 ThrMetAlaGlnMetValIleSerAsnIlePheAsnThrAspTyrHisMetAs 167
 167 nMetThrHisPheTyrValAlaPheAlaAlaTyrPheGlyValAlaThrValAlaT 184
 504 CATGATGACATCTACGTTCGACGACCTATTTTGGGCTGTGTGGCCT 553
 554 GGTGGCTCGCAAGGCTCTACCCGAGGAGACGAGAGATAAAGATAGACA 603
 184 rPcysLeuProLysProLeuProAspIleLysGlnAspLysAspGlnIle 200

604 GCAAGATACCCAGTTTGTCGCCATGCTGGGGCGCCCTTCTTGATG 653
 201 AlaTrIleProSerLeuSerAlaMetLeuGlyTrIlePheLeuTrpMe 217
 654 GTTCTGGCAAGTTTCAACTGCTGCTGCTGAGAAGTCAATCGAAGA 703
 217 tPheTrProSerPheAsnSerAlaLeuLeuArgSerProIleGlnArgL 234
 704 AGAATGCCGTGTTCACACCTTACTATGCTGTAGACAGTACGGTGTGACA 753
 234 ySAsnAlaValAlaPheAsnThrTyrTyrAlaLeuAlaValSerValAlaThr 250
 754 GCCATCTAGGGTCTATCTGCTGCTCACCCCAAGGAGATGACGAAAGAC 803
 251 AlaIleSerValSerSerLeuAlaHisProGlnGlyLysIleAsnMetH 267
 804 TTATGTGCACAGTGGCGTGTGGCAGAGGCGTGGCTGGGTACTCTGT 853
 267 rTyrMetHisAsnAlaValLeuAlaGlyGlyAlaValAlaGlyThrSerC 284
 854 GTCACCTGATCCCTTCTGCGGTGGCTTGGCATGCTGGCTGTGGCT 903
 284 ySHsLeuIleThrSerProTrpLeuAlaMetValLeuGlyLeuValAla 300
 904 GGGCTGATCTCCGTGCGGGAGACCAAGTACCTGCCGGGTGTGTAACG 953
 301 GlyLeuIleSerIleGlyGlyAlaLysCysLeuProGlyCysCysAsnArg 317
 954 AGTGTGGGGATTCGCCACAGCTCCATCATGAGGCTACACTTCAGCTTGC 1003
 317 gValLeuGlyIleHisAspSerSerValMetHisTyrAsnPheSerLeuL 334
 1004 TGGCTGCTGTGGAGAGATCATCTATGCTGCTGCTGCTGCTGATACC 1053
 334 euGlyLeuLeuGlyGlyIleThrTyrIleValLeuMetValLeuIstHr 350
 1054 GTGGAGCGCGCAATGGCATGATTGGCTTCCAGAGCTCTCATCAATGG 1103
 351 ValGlyAlaGlyAsnGlyMetValGlyPheGlnValLeuValSerThrGl 367
 1104 GGAATCAGCTTGGCCATGATAGCTGTCAAGTGTGCTGTGCTGACAG 1153
 367 yGlnLeuSerLeuAlaLeuAlaIleAlaValThrSerLysLeuLeuThrG 384
 1154 GTTGTCTCTTAATCTTAAATATGGAAGACCCGATGAGGCTTAATAT 1203
 384 lYleuLeuLeuAsnLeuLysIleTrpLysAlaProHisAlaAlaLysTyr 400
 1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
 401 PheAspAspGlnValPheTrpLysPheProHisLeuAlaValGlyPhe 416
 seq_name: SwissProt_39:RHL_MACMU
 seq_documentation_block:
 ID RHL_MACMU STANDARD; PRT; 416 AA.
 AC 028849;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBL_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE= Bone marrow;
 RX MEDLINE=94223693; PubMed=7909570;
 RA Mouro I., Le van Kim C., Cherif-Zahar B., Salvignol I., Blancher A.,
 Carton J.-P., Colin Y.;

RT "Molecular characterization of the Rh-like locus and gene transcripts
from the rhesus monkey (Macaca mulatta).";
RL J. Mol. Evol. 38:169-176(1994).
CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: S70343: AAB30637.1; -.
DR InterPro: IPR001905; -.
DR InterPro: IPR002229; -.
DR Pfam: PF00909; Ammonium_transp; 1.
DR PRINTS: PR00342; RHESUSRHD.
KM Erythrocyte; Transmembrane; Polymorphism.
FT INIT MET 0 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT VARIANT 1 55 MISSING (IN MAC-B).
FT VARIANT 70 70 A -> K (IN MAC-B).
FT VARIANT 106 106 A -> V (IN MAC-B).
FT VARIANT 191 191 P -> L (IN MAC-B).
FT VARIANT 261 261 G -> R (IN MAC-B).
FT VARIANT 342 342 Y -> C (IN MAC-B).
SQ SEQUENCE 416 AA; 45672 MW; 3BFD939BEA79AF6F CRC64;

alignment_scores: Quality: 1642.00 Length: 416
Ratio: 4.265 Gaps: 0
Percent Similarity: 92.548 Percent Identity: 77.404

alignment_block:

US-09-600-714-41 x RHL_MACMU ..

Align seg 1/1 to: RHL_MACMU from: 1 to: 416

```

4 AGCTTAAAGTACCGCGGTGTGTCGCGGCGCTGCGCCCTCTGGGCGCT 53
1 SerSerLysTYrProArgSerValAlaArgCysLeuProLeuTrpAlaLe 17
54 AACCTGGAACGAGCTCATCTTCCTCTCTATTTTAAACCCCTATNG 103
17 urThrLeuAlaAlaLeuIleLeuLeuPhePheThrTrpTYrA 34
104 ACGCTTCCTAGAGATCAAAAGGGCTGCTGCGATCCTATCAAGTTGGC 153
34 spAlaSerLeuGluAspGlnIlysglyLeuValAlaSerTYrGlnValLys 50
154 CAAGATCTGACCGGTGATGGCGCATTTGGCTTGGCTCTACCTGAG 203
51 GlnAspLeuThrValMetAlaValIleuGlyLeuGlyPhePheThrSerAs 67
204 TTTCGGGAGACACAGCTGAGACAGTGTGCTCAACTCTTCATGAGTNG 253
67 nleuArgIAsnSerItrPserSerValAlaPheAsnleuPheLeuA 84

```

```

254 CGCTTGTGTGCAGTGGGCAATCCTGCTGAGCGGCTTCTGACCACTTC 303
84 IaleuGlyValGlnTrpAlaIleLeuLeuAspGlyPheLeuSerGlnPhe 100
304 CCTTCTGGGAAGGTGCATCACTGTCAGTATTCGTCGGCGCACCAT 353
101 SerProGlyIysValAlaIleLysLeuPheSerIleAspLeuAlaThrAr 117
354 GAGTGCCTTGTGGTGTGATCTCATGATGGATGATGCTGTCTTGGGGAAGTCA 403
117 gSerThrIleSerMetLeuIleSerMetAsnAlaValLeuGlyValA 134
404 ACTTGGCGCACTGTGTGTGATGTGCTGTGTGAGGTACAGCTTTAGGC 453
134 snLeuValGlnLeuValAlaMetGluLeuValGluLeuThrValPheGly 150
454 AACCTGAGGATGTCATCAAGTAAATATCTTCAACACAGCTCCACATGAA 503
151 ThrMetArgIleValIleAsnAsnIlePheLysIleAspTYrGlyMetAs 167
504 CATGATGCACATCTACGTGTTCGACGCTATTTTGGGCTGTGTCGCGCT 553
167 nMetMetIleIleHisValPheAlaIleAlaTYrPheGlyLeuThrValAlaIat 184
554 GGTGCTGTCCAAAGCCTCTACCCGAGGAGGAGGAGATTAAGATCAGACA 603
184 rPcysLeuProLysProLeuProLysGlyThrGluAspLysTYrGlnThr 200
604 GCACGATACCCAGTTTGTCTGCCATGCTGGCGGCGCTTCTTGTGAT 653
201 ThrThrSerProSerLeuPheAlaMetLeuGlyThrLeuPheLeuTrpMe 217
654 GTTCTGGCGCAAGTTTCAACTGCTGCTGTGAGAGAGATTCATTCGAAAGA 703
217 tPheTrpProThrPheAsnSerIleAlaLeuLeuAsnProIleGluArgL 234
704 AGAATGCGGTGTCAACACCTACTATGCTGACGATCAGCGGTGTGACA 753
234 yAsnAlaValPheSerThrTYrAlaLeuAlaValSerAlaValThr 250
754 GCCATCTCAGGCTCATCTTGGCTCACCCTCCAAAGGAGGATCAGACAGAC 803
251 AlaIleSerValSerSerLeuAlaHisProGlyLysIleAsnMetTh 267
804 TTATGTGCACAGTGGGTTGGTGGCGAGGCGGTGCTGGGTACTCGCT 853
267 rTYrMetHisAsnAlaAlaLeuAlaGlyIlyValAlaLeuSerAlaSerC 284
854 GTACACCTGATCCCTTCTCCGCTTGGCTTGGCATGCTGCTGGTGTGCGCT 903
284 yHisValIleHisSerProTrpIleAlaMetValLeuGlyLeuValAla 300
904 GGGCTGATCTCGTGGGAGGAGCAAGTACCTGCCGGGTGTGTGAACG 953
301 GlyLeuIleSerIleGlyIlyAlaLysCysLeuProValLysPheAsnAr 317
954 AGTGGGGGATTCGCCACAGTCATCATGAGGCTGAGCAACTCAGCTTGC 1003
317 gValLeuGlyIleHisGluSerHisSerValHisTYrThrPheGlyLeuP 334
1004 TGGGTCTCTTGGAGAGATCATCTTACATTTGTCTGCTGGTGGTATGAC 1053
334 roAlaLeuLeuGlyIlyIleThrTYrIleValIleMetAlaLeuAlaArgL 350
1054 GTCGGAGCGCGCAATGCATGATTTGGCTTCCAGGTCTTCCTCAGATTGG 1103
351 ValTrpAlaSerSerAsnMetIleGlyPheGlnValLeuLeuSerThG1 367
1104 GGAATCAGCTTGGCCATGATAGCTTCACGCTGCTGCTCCCTGACAG 1153
367 yThrLeuSerIleuAlaMetAlaMetSerIleThrSerIlyLeuLeuTrng 384
1154 GTTGTCTCTTAATCTTAATAATATGGAAGACCTCATGAGGCTAAATAT 1203

```

|||||
 384 LysLeuLeuAsnLeuLysIleTrpLysGlyProHisValAlaLysTyr 400

1204 TTGATGACCAAGTTTCTGAGAGTTCCATTTGGCTGTGGATT 1251
 |||||
 401 PheAspArgGlnAlaPheTrpGlnPheProHisLeuAlaGlyPhe 416

seq_name: SwissProt_39:RHL_MACFA

seq_documentation_block:

ID RHL_MACFA STANDARD; PRT; 416 AA.

AC 028481;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=bone marrow;

RA MEDLINE=95085595; PubMed-7993375;

RA Salviagnol I., Blancher A., Calvas P., Clayton J., Socha W.W.,

RA Collin Y., Ruffie J.;

RT "Molecular genetics of chimpanzee Rh-related genes: their

RT relationship with the R-C-E-F blood group system, the chimpanzee

RT counterpart of the human rhesus system.";

RL Biochem. Genet. 32:201-221(1994).

CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: L37054; AAA65628.1; -

DR InterPro: IPR001905; -

DR Pfam: PF00909; Ammonium_transp. 1.

DR PRINTS: PR00342; RHESUSRHD.

KW Erythrocyte; Transmembrane.

FT INIT_MET 0 BY SIMILARITY.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 43 63 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 124 144 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 237 257 POTENTIAL.

FT TRANSMEM 264 284 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 330 350 POTENTIAL.

FT TRANSMEM 357 377 POTENTIAL.

SO SEQUENCE 416 AA; 45877 MW; 17832A58D2299A9 CRC64;

alignment_scores:

Quality: 1634.00

Ratio: 4.311

Percent Similarity: 91.106

Percent Identity: 78.125

alignment_block:

US-09-600-714-41 x RHL_MACFA ..

Align seg 1/1 to: RHL_MACFA from: 1 to: 416

4 AGCTCTAGTACCGCGGCTGTCGGCGCTGCGCCCTGCGGCGCT 53
 1 SerSerLysTyrProHisValArgCysLysLeuProLeuTrpAlaLe 17
 54 AACACTGGAAACAGCTCTCATTTCTCTTCTATTTTACCACATAG 103
 17 uThrLeuGlnAlaLeuLeuLeuPhePhePheThrTyrTyrTyr 34
 104 ACGCTTCTTAGAGATCAAAAGGCGCTGCGCATCTATCAAGTTGC 153
 34 spAlaSerLeuGlnAlaSpGlnLysLeuAlaSerTyrGlnValCys 50
 154 CAAGATCTGACCGTGTATGGCGGCGCATGGCTTGGCTTCCACCTGAG 203
 51 GlnAspLeuThrValMetAlaValLeuGlyLeuGlyPhePheThrSer 67
 204 TTTCGGAGACACAGCTGAGACAGTGGCCTTACCTTTCATGCTGC 253
 67 nLeuAlaArgAsnSerTrpSerSerValAlaPheAsnLeuPheLeu 84
 254 CGCTTGGTGTGACATGGGCAATCTGCTGGACGCGCTTCTGACCATC 303
 84 lAlaGlyValGlnTrpAlaLeuLeuAspGlyPheLeuSerGlnPhe 100
 304 CCTTCTGGAGAGTGTGCATCATCACTGCTTCAGTATTCGGCTGCGCACCAT 353
 101 SerProGlyLysValAlaLeuLysLeuPheSerLeuArgLeuAlaTrp 117
 354 GAGTGTCTTGTGCTGCTGATCTCAGTGGATGCTGTGGGAGAGTCA 403
 117 gSerThrThrSerMetLeuIleSerMetAsnAlaValLeuGlyVal 134
 404 ACTTGGGCGCATGGTGTGATGCTGCTGGGAGGAGACAGCTTTAGGC 453
 134 snLeuAlaGlnLeuValValMetGlnLeuValGlnLeuThrValPheGly 150
 454 AACCTGAGAGATGATCATCATATATATCTTCAACACAGACTCCACATGAA 503
 151 ThrMetArgLeuValIleTyrAsnIlePheLysTyrGlyMetAs 167
 504 CATATGACATCTACAGTTCGCGACCGCTATTTGGCGTGTGCGCT 553
 167 nMetMetHisIleHisValPheAlaAlaLysPheGlyLeuThrValAla 184
 554 GGTGCGCGCAAGCGCTCTACCGAGGAGGAGGATTAAGATCAGACA 603
 184 TrpCysLeuProLysProLeuProLysGlyTrpGlnAspLysTyrGlnThr 200
 604 GCAACGATACCCAGTTGTGTGCCATGCTGGCGCGCTCTTCTTGGAT 653
 201 ThrThrSerProSerLeuPheAlaMetLeuGlyThrLeuPheLeuTrp 217
 654 GTTCTGGCAAGTTTCAACTGCTGCTGCTGAGAAAGTCAATGCAAGA 703
 217 tPheTrpProThrPheAsnSerAlaLeuLeuAsnProIleGlnArgL 234
 704 AGAATGCGGTGTTCAACACACTATGCTAGACAGGAGGCTGTGACA 753
 234 yAsnAlaValAlaPheSerThrTyrTrpAlaLeuAlaValSerAlaValThr 250
 754 GCCATCTCAGGTCATCTTGGCTCACCCCAAGGAGATCAGACAGAC 803
 251 AlaIleSerValSerSerLeuAlaHisProGlnArgLysIleAsnMet 267
 804 TTATGTACACAGTGGTGTGGCAGGAGGCGTGTGGGATACCTGCT 853
 267 rTyrMetProAsnAlaGlyLeuAlaGlyGlyValAlaValAlaSerC 284
 854 GTCACTGATCCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
 284 yShiValIleHisSerProTrpIleAlaMetValLeuGlyLeuValAla 300

904 GGCGTGAATCTCCGCGGGGAGCCCAAGTACCTGCCGGGTGTGTAACCG 953
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 301 G1yleuileSerPheGlyAlaIalysCysleuProValCysPheAsnAr 317
 954 AGTCTGGGGATTCCCAAGCTCCATCATGGGCTACACTTACGCTTGC 1003
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 317 GValleuGlyIleHisGlySerHisSerMetHisIstYrThrPheGlyLeuP 334
 1004 TGGGTCTGCTGGAGAGATCATCTACATTTGGCTGGCTGGCTGATAC 1053
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 334 roAlaIleuLeuGlyIleIleThrYrIleValIleuMetAlaIleuArgVal 350
 1054 GTCCGAGCCGCAATGAGCATGATGGCTTCAGGCTCCCTCCACATTTGG 1103
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 351 PheThrAlaSerSerAsnMetIleGlyPheGlnValIleuLeuSerThrG 367
 1104 GGAACCTAGCTTGGCCATCTGATATGCTCTGACGTCTGGTCTCCAGAG 1153
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 367 YThrIleuSerIleuAlaMetAlaMetSerIleThrSerGlyLeuLeuThrG 384
 1154 GTTTCCTCCATAATCTTAATATGAAGACCTCATGAGGCTAAATAT 1203
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 384 IyleuLeuLeuAsnIleuYsIleThrPlyGlyProHisValAlaIalysYr 400
 1204 TTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTTGGATT 1251
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 401 PheAspAspGlnAlaPheThrGluPheProHisIleuAlaValGlyPhe 416

seq_name: SwissProt_39:RHL_HYLIPI

seq_documentation_block:
 ID RHL_HYLIPI STANDARD; PRT; 353 AA.

AC 028446;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last annotation update)
 DE RH-LIKE PROTEIN (RHESUS-LIKE PROTEIN).
 OS Hylobates pileatus (Pileated gibbon).
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates;
 OC NCBI_TaxID=9589;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=95085595; PubMed=7993375;
 RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
 RA Colin Y., Ruffie J.;
 RT "Molecular genetics of chimpanzee Rh-related genes: their
 RT relationship with the R-C-E-F blood group system, the chimpanzee
 RT counterpart of the human rhesus system.";
 RL Biochem. Genet. 32:201-221(1994).
 CC - FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
 HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC - SIMILARITY: BELONGS TO THE RH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L37051; AAA6625.1; -
 DR InterPro: IPR001905; -
 DR Pfam: PF00909; Ammonium_transp. 1.
 KM Erythrocyte; Transmembrane.
 FT INIT MET 0 BY SIMILARITY.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.

FT TRANSMEM 124 144 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 237 257 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 SQ SEQUENCE 353 AA; 38417 MM; 3CFD25B38033AB6A CRC64;

alignment_scores:
 Quality: 1527.00 Length: 398
 Ratio: 4.572 Gaps: 3
 Percent Similarity: 83.920 Percent Identity: 76.884

alignment_block:

us-09-600-714-41 x RHL_HYLIPI ..

Align seg 1/1 to: RHL_HYLIPI from: 1 to: 353

4 AGCTCTAGTACCGCGGCTGTGCTGGCGGCTGCCCTCGCGGCGCT 53
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1 SerSerIstYrProArgSerValArgGlyCysLeuProLeuThrAlaIe 17
 54 AACACTGGAACAGCTCTCATCTTCTCTCTATTTTATCCCATATG 103
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 17 uThrIleuGlnAlaIleuIleLeuIleuPhePhePheThrHisIstYrA 34
 104 ACGCTTCCCTTAGAGATCAAAAGGGCTCGTGGCATCCTATCAAGTTGGC 153
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 34 sPAlaSerIleuGlnAspGlnIysGlyIleuValAlaIleThrGlnValGly 50
 154 CAAGATCTGACCGGTGATGCGGCATTTGGCTTGGCTCTACCTCGAG 203
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 51 GlnAspLeuThrValMetAlaAlaIleuGlyIleuGlyPheLeuThrSerAs 67
 204 TTTCCGAGACACAGCTGAGCATGCTGGCTTCACCTCTTCATCTGCG 253
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 67 nLeuArgThrHisSerIstYrPheSerValAlaPheAsnIleuPheMetLeuA 84
 254 CGCTTGGTGTACAGTGGGCAATCTGTGTGACGCGCTTCGAGCCAGTTC 303
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 84 IAlaIleuGlyValGlnThrPAlaIleuLeuAspGlyPheLeuSerGlnPhe 100
 304 CTTTCTGGGAAGTGTGATCATCACACTGTTCAATTCGGCTGGCCACAT 353
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 101 ProProGlyIstValValIleIleIstLeuLeuSerIleArgLeuAlaIle 117
 354 GAGTCTTGTTCGCGTGTGATCTGATGATGATGATGATGATGATGATGAT 403
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 117 tSerAlaMetSerSerIleIleSerValGlyAlaValAlaIleuGlyIstValA 134
 404 ACTTGGCGCAGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 453
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 134 snIleuValGlnIleuValValMetValIleuValGlnIleThrAlaPheGly 150
 454 AACCTGAGGATGTGATCATGTAATATCTTCAACACAGCATCCACATGAA 503
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 151 ThrMetArgMetValIleAsnAsnIlePheIstYrHisIstValAs 167
 504 CATGATGACATCTTACGTTGCGAGCTATTTGGCTGCTGGTGGCT 553
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 167 nIleMetHisIleHisValPheAlaIleIstYrPheGlyLeuThrIstValA 184
 554 GGTGCTCTCCAAAGCTCTACCCGAGGAGGAGGAGGATAAAGATCAGACA 603
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 184 rPcYsLeuProIstYrProProProGlnGlyThrGlnAspIstGlnIle 200
 604 GCAAGATACCAAGTTTGTGTGCCATGTGCGGCGCTCTTCTTGAT 653
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 201 AlaAsnSerProSerIstYrAlaMetIleuGlyAlaIleuPheLeuThrP 217
 654 GTTCTGGCAAGTTTCAACTGCTGCTGCTGAGGATCCATTCGAAAGA 703
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 217 ePheThrProSerPheAsnSerAlaIleuLeuThrAsnProIleGlnArgL 234

```

704 AGATGCCGTTGTTCAACACCTACTATGCTGTAGCAGTCAGCTGTGACCA 753
|||||
234 ysaasnlavallphehnsnhrtyrtyralaleualavalsethrvalthr 250
|||||
754 GCCATCTCAGGGTCATCTCTGGCTCACCCCCCAAGGGAAGATCAGCAAGAC 803
|||||
231 AlaileSerValserSerleualahisProginglylsileAsnmetLth 267
804 TTATGTGACAGTGGGCTGTGGCGAGGAGGCTGGCTGGTACTCTGCT 853
|||||
267 rtyrmetHisasnalaValleualalelyValalavalalyltmtserC 284
|||||
854 GTCACCTATCCCTCTCCCTGGCTTGCATGCTGCTGGCTGTGGTGGCT 903
|||||
284 yshIsleuileSerSerProtrileuametalvalleuvalale 300
904 GGGCTGATCTCCGTCGGGGGAGACCAAGTACCTGCCGGGGTGTGAACCG 953
|||||
301 GlyleuileSerileglylylValalys.....CysLeu..Pro 312
954 AGTGTGGGGATGCCACAGCTCATGATGGGCTACAACTTCAGCTTGC 1003
312 .....
312 .....
1004 TGGGTCTGCTTGGAGAGATCATCTACATTGTCTGCTGCTGTGATACC 1053
312 .....
1054 GTCGGAGCCGCGCATGCGATGATGGCTTCAGCTCTCAGCATTTGG 1103
|||||
313 .....AsptleuProaspProleuGlnhistrrp 322
1104 GGAACCTAGCTTGGCATCGTATGATCTTCACGCTGCTCTCCGACAG 1153
|||||
323 GlytrmGlnleuGlyHisGlyAspserSerHisValtrpPheProaspPar 339
1154 GTTGGCTCTTAATCTTAATATGGAAGACCTCATGAGGC 1196
|||||
339 gPheAlarProlysserGlnasnmelGlnserTherCysGly 353

seq_name: SwissProt_39:RHAG_HUMAN
seq_documentation_block:
ID RHAG_HUMAN STANDARD: PRT: 409 AA.
AC 002094:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RHESUS BLOOD GROUP-ASSOCIATED GLYCOPROTEIN (ERYTHROCYTE PLASMA
DE MEMBRANE 50 KDA GLYCOPROTEIN) (RH50A).
GN RHAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow, and Liver;
RX MEDLINE-93038558; PubMed-141776;
RA Ridgwell K., Spurr N.K., Laguda B., Macgeoch C., Avenet N.D.,
RA Tanner M.J.;
RT Isolation of cDNA clones for a 50 kDa glycoprotein of the human
RT erythrocyte membrane associated with Rh (Ihesus) blood-group antigen
RT expression.
RT Biochem. J. 287:223-228(1992).
RL
[2]
RP SEQUENCE OF 1-39.
RX MEDLINE-89134163; PubMed-3146980;
RA Avenet N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J.,
RA Kumpel B.;
RT Protein-sequence studies on Rh-related polypeptides suggest the
RT presence of at least two groups of proteins which associate in the

```

```

RT human red-cell membrane.
RL Biochem. J. 256:1043-1046(1988).
RN [3]
RP VARIANT HEMOLYTIC ANEMIA ASN-79.
RX MEDLINE-96154189; PubMed-8563755;
RA Cherif-Zahar B., Raynal V., Gane P., Mattei M.-G., Bailly P.,
RA Gibbs B., Colin V., Cartron J.-P.;
RT "Candidate gene acting as a suppressor of the RH locus in most cases
RT of Rh-deficiency";
RT Nat. Genet. 12:168-173(1996).
CC -1- FUNCTION: ASSOCIATED WITH RHESUS BLOOD GROUP ANTIGEN EXPRESSION.
CC MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO HAVE A
CC TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ERYTHROCYTES.
CC -1- DISEASE: DEFECTS IN RHAG ARE A CAUSE OF A FORM OF CHRONIC
CC HEMOLYTIC ANEMIA ASSOCIATED WITH STOMATOCYTOSIS, AND
CC SPHEROCYTOSIS, REDUCED OSMOTIC FRAGILITY, AND INCREASED CATION
CC PERMEABILITY.
CC -1- SIMILARITY: BELONGS TO THE RH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X64594; CAA45883.1; -.
CC DR PIR: S29124; S29124.
CC DR MIM: 180297; -.
CC DR MIM: 268150; -.
CC DR InterPro: IPR001905; -.
CC DR InterPro: IPR002229; -.
CC DR Pfam: PF00909; Ammonium_transp; 1.
CC DR PRINTS: PR00342; RHESUSRD.
CC KW Erythrocyte; Glycoprotein; Transmembrane; Disease mutation.
CC FT DOMAIN 1 4
CC FT TRANSMEM 5 24
CC FT DOMAIN 25 56
CC FT TRANSMEM 57 72
CC FT DOMAIN 73 83
CC FT TRANSMEM 84 101
CC FT DOMAIN 102 119
CC FT TRANSMEM 120 134
CC FT DOMAIN 135 140
CC FT TRANSMEM 141 153
CC FT DOMAIN 154 175
CC FT TRANSMEM 176 190
CC FT DOMAIN 191 208
CC FT TRANSMEM 209 225
CC FT TRANSMEM 226 238
CC FT TRANSMEM 239 261
CC FT DOMAIN 262 274
CC FT TRANSMEM 275 285
CC FT TRANSMEM 286 292
CC FT TRANSMEM 293 311
CC FT DOMAIN 312 333
CC FT TRANSMEM 334 351
CC FT DOMAIN 352 371
CC FT TRANSMEM 372 383
CC FT DOMAIN 384 409
CC FT CARBOHYD 37 37
CC FT VARIANT 79 79
CC FT CONFLICT 2 2
CC FT CONFLICT 37 37
CC FT SEQUENCE 409 AA; 44199 MW; 1E5735FB18CB8E29 CRC64;

alignment_scores:
Quality: 628.50
Length: 416

```

Ratio: 2.131 Gaps: 9
Percent Similarity: 70.913 Percent Identity: 36.298

alignment block:

US-09-600-714-41 x RHAG_HUMAN ..

Align seg 1/1 to: RHAG_HUMAN from: 1 to: 409

```

25  GTGGGGGCGCTGCTGGCCCTGGGCGCCCTACACGTGAAGGAGCTTCAT 74
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1  MetArgPheThrPheProLeuMetAlaIleValLeuGluIleAlaMetCil 17
75  TCTGCTCTTATTTTATTTTACCACATAT.....GACGCTTCCTTAGAG 118
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
17  eValLeuPheGlyLeuPheValIgluTyrgluThrAspGlnThrValLeuG 34
119  ATCAAAAGGGGCTGCGCATCCTATCAAGTTGGC..... 153
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
34  IugIleuAsnIleThrLysProThrAspMetGlyIlePhePheGluLeu 50
154  .....CAAGATCTGACGCGGCGCCATTTGGCTTGCGGCTT 191
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
51  TyrProLeuPheGlnAspValAlaHisValMetIlePheValIgluPheGly 67
192  CCGACCTCGAGTTTCCGAGACACAGCTGAGAGAGTGGCGCTTCAAC 241
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
67  eLeuMetThrPheLeuLysLysTyrglyPheSerSerValIgluIleAsn 84
242  TCTTCATGCTGGCGCTTGTGTCGAGTGGGCAATCCTGCTGAGCGCTTC 291
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
84  euleValAlaIleValLeuGluIleGlnTygIleValIleGlnIle 100
292  CTGAGCCAGTTCCTTCTGGGAGAGTGGTCATCACATGTCAGATTGCG 341
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
101  Leu...GlnSerGlnGlyGlnLysPheAsnIleGlyIleLysAsnMetCil 116
342  GCTGGCCACCATGAGTGGCTTGTGCGGCTGCTGATCTCACTGATGCTGCT 391
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
116  eAsnAlaAspSerAlaIleAlaThrValLeuIleSerPheGlyAlaVal 133
392  TGGGGAAGTCACTTGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 441
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
133  euGlyLysThrSerProThrGlnMetLeuIleMetThrIleLeuIle 149
442  ACAGCTTAGGACCTGAGTGGTGCATGCTATATATCTTCAACACAGA 491
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
150  ValPhePheAlaHisAsnGluTyrlleuValSerGluIlePheLysAla 166
492  CTACACATGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 541
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
166  RasPleGlyAlaSerMetThrIleHisAlaPheGlyAlaTyrlle 183
542  TGTCTGTGGCTGGTGGCTGGCCAAAGCCT...CTACCCGAGGAGGAG 588
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
183  euaIleValAlaGlyIleLeuTyrgArgSerGlyLeuArgLysIleGlu 199
589  GATAAAGATGACAGACAGACATACCCAGTTTGTGCTGCCATCTGGGCG 638
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
200  Asn...GluGlnSerAlaIleTyrlleSerAspLeuPheAlaMetIleGly 215
639  CCTCTTCTGTGGATGTTCTGGCCAAAGTTTCACTGCTGCTGCTGAGAA 688
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
215  rLeuPheLeuThrMetPheThrProSerPheAsnSerAlaIleAlaGlu 232
689  GTCCAAATGAAAGAGAAATGCGGTTCACACACTACTATGCTTAGCA 738
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
232  roGlyAspLysGlnCysArgAlaIleValAspThrTyrllePheSerLeu 248
739  GTACGCGGTGATGACCATCTCAGGATCATCTGGTGGTGGTGGTGGTGG 788
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
249  AlacysValleuThrAlaPheAlaPheSerSerLeuValIgluHisArg 265
789  GAAGATCAGCAGAACTTATGTGCACAGTGGCGGTGTGGCAGGAGCGTGG 838

```

```

|||||.....  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
265  yLysLeuAsnMetValHisIleGlnAsnAlaThrLeuAlaGlyAla 282
839  CTGTGGGTACCTGCTGTCACCTGATCCCTTCTCGGTGGCTGACGATG 888
    |||||.....  :::::  :::::  :::::  :::::  :::::  :::::  :::::
282  IValIgluThrCysAlaAspMetAlaIleHisProPheGlySerMetCil 298
889  CTGGATCTTGTGCGCTGATCTCGCTGCGGAGGAGGAGGAGGAGGAGG 938
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
299  IlegIleSerIleAlaGluMetValSerValLeuGlyTyrllePheLeu 315
939  GGGGTGTGTACCCAGATGCTGGGAGATTCCACACAGCTCCATCATGG 988
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
315  rProLeuPheThrThrLysLeuArgIle...HisAspThrCysGlyVal 331
989  ACAACTTC...AGCTTGTGGTCTGCTGGGAGAGAGATCATCATGTTG 1035
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
331  IsAsnLeuHisGlyLeuProGlyValIgluGlyLeuAlaGlyIleVal 347
1036  CTGCTGTGCTTGTATACCGTCGAGCGGCAATGAGCATGATGGCTTCCA 1085
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
348  AlValAlaMet.....GlyAlaSerAsnThrSerMetAlaMetCil 361
1086  GGTCTCTCTCAGCATTTGGGAGACTGAGCTTGGCCATGCTGATGCTCA 1135
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
361  nAla.....AlaAlaLeuGlySerSerIleGlyThrAlaVal 374
1136  CGTGTGCTCTCTGACAGTTTGTGCTTAAATCTTAAATGGAAGCA 1185
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
374  AlGlyGlyLeuMetThrGlyLeuIleLeuLysLeuProLeuThrGlyGln 390
1186  CCTCATGAGGCTTAAATATTTTATGATGACCAAGTTTCTGGAAGTTCC 1233
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
391  ProSerAspGlnAsnCysTyrlleAspSerValTyrllePheValPro 406

seq_name: SwissProt_39:SRCH_HUMAN

seq_documentation_block:
ID  SRCH_HUMAN  STANDARD;  PRT;  699 AA.
AC  P23327;
DT  01-NOV-1991 (Rel. 20, Created)
DT  01-NOV-1991 (Rel. 20, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BINDING PROTEIN
DE  PRECURSOR.
GN  HRC OR HCP.
OS  Homo sapiens (human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Skeletal muscle;
RX  MEDLINE=91244309; Pubmed=2037293;
RA  Hofmann S.L., Topham M., Hsieh C.L., Francke U.;
RT  cDNA and genomic cloning of HRC, a human sarcoplasmic reticulum
RT  protein, and localization of the gene to human chromosome 19 and
RT  mouse chromosome 7.
RL  Genomics 9:656-669(1991).
CC  -!- FUNCTION: HCP MAY PLAY A ROLE IN THE REGULATION OF CA(2+)
CC  SEQUESTRATION OR RELEASE IN THE SR OF SKELETAL AND CARDIAC
CC  MUSCLE.
CC  -!- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM LUMEN.
CC  -!- SIMILARITY: STRONG, TO RABBIT HRC.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.sib.ch/announce/
CC  or send an email to license@sib.ch).

```


OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 644 to 924 of the genome.";
 RL DNA Res. 2:153-166(1995).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE APT/MEP/NRGA FAMILY OF AMMONIUM
 CC TRANSPORTERS (TC 2.49).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, D64004; BAAL0631.1; -;
 DR InterPro: IPR001905; -;
 DR Pfam: PF00909; Ammonium_transp; 1.
 DR PROSITE: PS01219; AMMONIUM_TRANS; 1.
 KM Hypothetical protein: Transport: Transmembrane.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 178 198 POTENTIAL.
 FT TRANSMEM 211 231 POTENTIAL.
 FT TRANSMEM 245 265 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 FT TRANSMEM 315 335 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 370 390 POTENTIAL.
 FT TRANSMEM 399 419 POTENTIAL.
 FT TRANSMEM 438 458 POTENTIAL.
 SQ SEQUENCE 507 AA; 53580 MW; 1840AE99F8A08BE6A CRC64;

alignment_scores:
 Quality: 138.50 Length: 369
 Ratio: 0.710 Gaps: 18
 Percent Similarity: 52.846 Percent Identity: 23.306

alignment_block:
 US-09-600-714-41 x Y108_SYNY3 ..

Align seg 1/1 to: Y108_SYNY3 from: 1 to: 507

```

19 CGGTCGTCCGGCGC.....TGCCTGCCCT 44
   ||| |||||
34 ArgpHeValArgylsleasnsErProtrpleuAlaCysValProle 50
   ||| |||||
45 CTGGGCGCTTAACATG.....GAAGCAGCTCATTCCTCTCTCT 85
   ||| |||||
50 uThrAlaLeuValAlaIletrpAsnAlaIleAlaIleAlaIleAspT 67
   ||| |||||
86 ATTTTATACCACTATGACGCTTCCTAGAGATCAAAAGGCGCTG 135
   ||| |||||
67 hrGlulIleVal.....AsnIleThrValIleGluThrValAsnGluAsnVal 81
   ||| |||||
136 GCATCTATCA.....GTGGCCACA 158
   ||| |||||
82 AlaThrLeuGlnGlyThrLeuAsnAlaIletrpIleLeuIleAlaIle 98
   ||| |||||
159 TCTGACCGTATGGCGCCATTGGCTGGCTTCCTCACCTGACGATTC 207
   ||| |||||

```

```

98 eleuValIlePheMetAsnAlaGlyPheGlyMetLeuGluThrGlyLeu 115
208 .....CGGAGACACAGCTGAGACAGTGGCTTCAACCTTCATGCTG 252
   ||| |||||
115 ysArgGlnLysAsnAlaValAsnIleLeuThrLysAsnLeuIleValPhe 131
   ||| |||||
253 CGCGTTGGT.....GTCCAGTGGGCAATCCTGCTGCAGCGCTTC 291
   ||| |||||
132 AlaLeuAlaThrIleAlaIletrpAlaIle.....GlyPheSerIle 145
   ||| |||||
291 .....
145 uMetPheGlySerSerGlyAsnProPheValGlyPheGlyIlePheIle 162
   ||| |||||
292 .....CTGAGCCAGTTCCTCTGGGAG 315
   ||| |||||
162 euserGlyAspHisThrAsnGlyLeuSerProPheProGlnGlyLeu 178
   ||| |||||
316 GTGGTCATCACACCTGTTAGTATTCGGCTGCCACACATGACTGTTGTC 365
   ||| |||||
179 ProValAlaValIlePheLeuPheGlnValAla...PheSerAlaThrAl 194
   ||| |||||
366 GGTGCTGATCTCAGTGCATGCTGCTGGGAGGTCACATGGGCGAGT 415
   ||| |||||
194 AlaThrIleValSerGlyAlaValAlaGluArgIleLysPheAsnGluP 211
   ||| |||||
416 TGGTGGTATGATGCTGCTGGAGGTGACAGCTTACGCCACCTGAGATG 465
   ||| |||||
211 heLeuIlePheSerValLeu.....LeuValGlyIleAlaIleArg 223
   ||| |||||
466 GTCATCATTAATATCTTCAACACAGC.....TACCACAT 500
   ||| |||||
224 ProIleThrGlyHisThrValIletrpAspAlaGlyLeuTrpLeuTythr 240
   ||| |||||
501 GAACATGATGCAC.....ATCAGCTGTTGGAGGCT 532
   ||| |||||
240 GlYpHeMetAspPheAlaGlySerThrValValHisSerValGlyGly 257
   ||| |||||
533 ATTTGGGCTGTCTGTGCTGTGCTGCCAAGCCTTACCCGAGGGA 582
   ||| |||||
257 rpAlaAlaLeuAlaGlyAlaPheLeuLeuGlyProArgLeuGlyLysPhe 273
   ||| |||||
583 ACGAGATTAAGATCAGACAGCAAGATPCCAGTTGTCT..... 624
   ||| |||||
274 ValAspGlyArg.....ProGlyAlaIleProGlyHisAsnMetGlyP 288
   ||| |||||
625 .GCCATGCTGGGCGCCCTTCTTGTGATGTCGCGCAGATTCACT 673
   ||| |||||
288 eAlaMetLeuGlyCysLeuIleLeuTrpIleGlyTrpPheGlyPheAsn 305
   ||| |||||
674 CTGCTGTGCTGAGAGTCCAAATCGAAGAGAAAGATGCCGTTCAAAC 723
   ||| |||||
305 roGly.....SerGlnLeuAlaAlaAspGlnAlaCys.....Ala 316
   ||| |||||
724 TACTATGCTTAGACAGTACGCTGTGACAGCCATCTGAGGTCATCTT 773
   ||| |||||
317 TyrIleAlaValThrThrAsnLeuAlaIleSerAlaGlyGlyLeuThrAl 333
   ||| |||||
774 GGCTCAC.....CCCAAGGAGATCAGCAAGATTATGTC 811
   ||| |||||
333 aThrHeThrSerTrpLeuLysAspGlyLysProAspLeuThrMetVal 350
   ||| |||||
812 ACAGTGGGCTGTTGGCAGAGCGCGGTGGGTGACCTGCTGTCACCTG 861
   ||| |||||
350 leAsnGlyValLeuAlaGlyLeuValGlyIleThrAlaGlyCysAlaGly 366
   ||| |||||
862 ATCCCTCTCCGCTGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGAT 911
   ||| |||||
367 Val...SerTyrTrpGlySerValIleIleGlyGlyIleAlaGlyIle 382
   ||| |||||
912 CTCGCTC 918
   ||| |||||
382 uValVal 384

```

```

seq_name: swissprot_39: FHUB_ECOLI
seq_documentation_block:
ID FHUB_ECOLI STANDARD: PRT: 660 AA.
AC P06712; P77372;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FERRICHRONE TRANSPORT SYSTEM PERMEASE PROTEIN FHUB (FERRICHRONE UPTAKE
DE PROTEIN FHUB).
GN FHUB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87014116; PubMed=3020380;
RA Koester W., Braun V.;
RT "Iron hydroxamate transport of Escherichia coli: nucleotide sequence
RT of the fhuc gene and identification of the protein.";
RL Mol. Gen. Genet. 204:435-442(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / W3110;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-23 FROM N.A.
RX MEDLINE=88038363; PubMed=2823072;
RA Burkhardt R., Braun V.;
RT "Nucleotide sequence of the fhuc and fhud genes involved in iron
RT ATP-binding proteins.";
RL Mol. Gen. Genet. 209:49-55(1987).
CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS INVOLVED IN THE HIGH-
CC AFFINITY TRANSPORT OF IRON(III)-FERRICHRONE INTO THE E.COLI
CC CELL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE FECD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04319; CAA27852.1; -.

```

```

DR EMBL: X04319; CAA27852.1; ALT_INIT.
DR EMBL: D26562; CAB20273.1; -.
DR EMBL: AE000124; AAC73264.1; -.
DR EMBL: U70214; AAB08583.1; -.
DR EMBL: X05810; CAA29236.1; -.
DR PIR: S07318; S07318.
DR PIR: S45222; S45222.
DR EcoGene: EG10303; fhuc.
DR InterPro: IPR000522; -.
DR Pfam: PF01032; Feecd_family; 2.
KW Transport; Iron transport; Transmembrane; Inner membrane.
FT TRANSMEM 5
FT TRANSMEM 25
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 118 138 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 197 217 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 447 467 POTENTIAL.
FT TRANSMEM 479 499 POTENTIAL.
FT TRANSMEM 528 548 POTENTIAL.
FT TRANSMEM 567 587 POTENTIAL.
FT TRANSMEM 607 627 POTENTIAL.
FT TRANSMEM 635 655 POTENTIAL.
FT TRANSMEM 655 655 L -> V (IN REF. 5).
FT CONFLICT 11 11 T -> I (IN REF. 1 AND 2).
FT CONFLICT 311 311 MISSING (IN REF. 1 AND 2).
FT CONFLICT 364 364
SQ SEQUENCE 660 AA; 70422 MW; 427313E7BA21B0BC CRC64;

alignment_scores:
Quality: 136.50 Length: 459
Ratio: 0.659 Gaps: 24
Percent Similarity: 45.098 Percent Identity: 21.351

alignment_block:
US-09-600-714-41 x FHUB_ECOLI ..
Align seg 1/1 to: FHUB_ECOLI from: 1 to: 660
127 GGGCTCGGCATCTTCAAGTGGCCAAAGATCTGACCGGATGGCGC 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 GYValAlaIhrGlyAlaGlnLeuGlyIleIhrValIhrThrLeuTPAl 110
177 CATGGCTTGGCTTCTCCTACCTGAGCTTCCGGAGACACAGCTGAGCA 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 allePro...GlyAlaMetAlaSerGlnPheAlaAlaGlnAlaGlyAlaC 126
227 GTTGGCTTACACCTCTCATGCTGGCGCTGTGGTGGCAGTGGCATC 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 yVal...ValGlyLeuIleValPheGlyValAlaIhrPGLys 139
277 CTGCTGAC... 285
|||||:
140 ArgLeuSerProValIhrLeuIleLeuAlaGlyLeuValValSerLeuTy 156
285 ..... 285
156 rcysGlyAlaIleAsnGlnLeuValIlePheHisHisAspGlnLeuG 173
286 .....GGCTTCTGACCGCAG..... 300
173 InsMetPheLeuTrpSerThrGlyIhrLeuIhrGlnIhrAspTrpGly 189
301 .....TTCCTCT.....GGAAAGTGGATCATCACACT 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
190 GlyValGlnArgLeuTrpProGlnLeuLeuGlyGlyValMetLeuThrLe 206

```


